SELECTION AND BREEDING PROGRAMS IN AQUACULTURE

Selection and Breeding Programs in Aquaculture

edited by

TRYGVE GJEDREM

AKVAFORSK, Institute of Aquaculture Research AS, Norway

A C.I.P. Catalogue record for this book is available from the Library of Congress.

ISBN-10 1-4020-3341-9 (HB) Springer Dordrecht, Berlin, Heidelberg, New York
ISBN-10 1-4020-3342-7 (e-book) Springer Dordrecht, Berlin, Heidelberg, New York
ISBN-13 978-1-4020-3341-4 (HB) Springer Dordrecht, Berlin, Heidelberg, New York
ISBN-13 978-1-4020-3342-1 (e-book) Springer Dordrecht, Berlin, Heidelberg, New York

Published by Springer,
P.O. Box 17, 3300 AA Dordrecht, The Netherlands.

Cover photo by Vidar Vassvik

Printed on acid-free paper

All Rights Reserved
© 2005 Springer
No part of this work may be reproduced, stored in a retrieval system, or transmitted
in any form or by any means, electronic, mechanical, photocopying, microfilming, recording
or otherwise, without written permission from the Publisher, with the exception
of any material supplied specifically for the purpose of being entered
and executed on a computer system, for exclusive use by the purchaser of the work.

Printed in the Netherlands.

PREFACE

Although aquaculture as a biological production system has a long history, systematic and efficient breeding programs to improve economically important traits in the farmed species have rarely been utilized until recently, except for salmonid species. This means that the majority of aquaculture production (more than 90 %) is based on genetically unimproved stocks. In farm animals the situation is vastly different: practically no terrestrial farm production is based on genetically unimproved and undomesticated populations. This difference between aquaculture and livestock production is in spite of the fact that the basic elements of breeding theory are the same for fish and shellfish as for farm animals. One possible reason for the difference is the complexity of reproductive biology in aquatic species, and special consideration needs to be taken in the design of breeding plans for these species. Since 1971 AKVAFORSK, has continuously carried out large scale breeding research projects with salmonid species, and during the latest 15 years also with a number of fresh water and marine species. Results from this work and the results from other institutions around the world have brought forward considerable knowledge, which make the development of efficient breeding programs feasible. The genetic improvement obtained in selection programs for fish and shellfish is remarkable and much higher than what has been achieved in terrestrial farm animals.

The main purpose of this book is to summarise quantitative genetic theory, the basic ingredient for selective breeding, and present methods for developing efficient breeding programs for aquatic species based on the latest scientific findings. Many books have been published about the development of breeding programs in farm animals, but so far no similar book has been available for aquatic species.

We hope that university students, biologists and consultants who plan and implement breeding programs in aquaculture will make use of the book. We also hope that the book will encourage the industry to use selective breeding in order to increase productivity and improve the utilisation of available feed, land and water resources for the benefit of the growing human population.

Many people have assisted in the writing process of this book. In particular, my colleagues at AKVAFORSK and in some other collaborating institutions have made crucial contributions. Their efforts, support and advice have been decisive for the completion of the work. I also owe a lot to Ms. Grethe Tuven (Department of Animal Science and Aquaculture, Agricultural University of Norway) who has been in charge of processing the figures, and to Dr. Ben Hayes and Dr. Nick Robinson (AKVAFORSK) who read the manuscripts and made the language readable and understandable. Finally I am grateful to the leaders of AKVAFORSK, who have encouraged us to complete and publish the book.

Trygve Gjedrem

Ås June 2004

CONTRIBUTORS

Øivind Andersen, Institute of Aquaculture Research, AKVAFORSK, P.O.BOX 5010, 1432 Ås, Norway

Hans Bernhard Bentsen, Institute of Aquaculture Research, AKVAFORSK, P.O.BOX 5010, 1432 Ås, Norway

Kjersti Turid Fjalestad, Norwegian Institute of Fisheries and Aquaculture Research, P.O.BOX 6122, 9291 Tromsø, Norway

Trygve Gjedrem, Institute of Aquaculture Research, AKVAFORSK, P.O.BOX 5010, 1432 Ås, Norway

Bjarne Gjerde, Institute of Aquaculture Research, AKVAFORSK, P.O.BOX 5010, 1432 Ås, Norway

Ben Hayes, Institute of Aquaculture Research, AKVAFORSK, P.O.BOX 5010, 1432 Ås, Norway

Kari Kolstad, Institute of Aquaculture Research, AKVAFORSK, P.O.BOX 5010, 1432 Ås, Norway

Theo H.E. Meuwissen, Department of Animal Science and Aquaculture, P.O.BOX 5003, 1432 Ås, Norway

Ingrid Olesen, Institute of Aquaculture Research, AKVAFORSK, P.O.BOX 5010, 1432 Ås, Norway

Terje Refstie, AKVAFORSK Genetics Center, 6600 Sunndalsøra, Norway

Morten Rye, AKVAFORSK Genetics Center, 6600 Sunndalsøra, Norway

Anna Sonesson, Institute of Aquaculture Research, AKVAFORSK, P.O.BOX 5010, 1432 Ås, Norway

Jørn Thodesen, AKVAFORSK Genetics Center, 6600 Sunndalsøra, Norway.

John A. Woolliams, Roslin Institute, Roslin, Midlothian EH25 9PS, United Kingdom

TABLE OF CONTENTS

Preface		v
Contributors		vii
1.	**Status and scope of aquaculture**	1
	Trygve Gjedrem	
1.1	Introduction	1
1.2	Domestication	2
1.3	History of animal breeding	3
1.4	Aim of breeding - Why change a population?	5
1.5	Effect of breeding programs in farm animals	5
1.5.1	What has been accomplished?	5
1.6	Breeding in aquaculture	6
1.7	Present status of breeding in aquaculture	7
2.	**Basic genetics**	9
	Trygve Gjedrem and Øivind Andersen	
2.1	Molecular genetics	9
2.1.1	DNA-molecule	9
2.1.2	Replication of DNA	14
2.1.3	Protein synthesis	14
2.2	Cell division	14
2.2.1	Mitosis	15
2.2.2	Meiosis	15
2.2.3	Location of genes	17
2.3	Mendelian inheritance	17
2.3.1	Example of mendelian traits in fish	18
2.4	Gene action	19
3.	**Population genetics**	23
	Øivind Andersen and Ben Hayes	
3.1	Introduction	23
3.2	Hardy-Weinberg equilibrium	23
3.3	Inbreeding	25
3.4	Crossbreeding	26
3.5	Selection	27
3.6	Mutation	28
3.7	Genetic drift	29
3.8	Migration	30
3.9	Genetic distance and population divergence	31
3.10	Summary	32

4.	**Polygenic inheritance**	35
	Hans Bernhard Bentsen	
4.1	Introduction	35
4.2	Limitations of the single gene model	35
4.3	A simplified polygenic model	37
4.4	How many genes?	40
4.5	Implications for the response to selection	40
4.6	Limitations of the simplified polygenic model	42
4.7	Summary	43
5.	**Basic statistical parameters**	45
	Trygve Gjedrem and Ingrid Olesen	
5.1	Variation in traits	45
5.1.1	Mean, standard deviation and variance	45
5.1.2	Coefficient of variation	47
5.2	Distribution of observations	47
5.2.1	Normal distributions	47
5.2.2	Non-normal distributions	49
5.3	Variance	51
5.3.1	Analysis of variance	51
5.3.2	Variance of a sum	53
5.3.3	Genetic variance	54
5.3.4	Environmental variance	54
5.3.5	Maternal variance	57
5.3.6	Genotype – environment interaction	58
5.4	Relationship between traits	59
5.4.1	Correlation and regression	59
5.4.2	Analysis of covariance	60
5.4.3	Genetic and phenotypic correlation	61
5.5	Heritability	64
5.6	Path coefficients	65
6.	**Kinship, relationship and inbreeding**	73
	Anna K. Sonesson, John A. Woolliams and Theo H.E. Meuwissen	
6.1	Introduction	73
6.2	Concepts	73
6.3	Calculating coefficients of kinship and relationship	76
6.3.1	Paths and loops	76
6.3.2	The tabular method	78
6.3:3	Properties of the relationship matrix (A matrix)	79
6.4	Consequences of inbreeding	81
6.4.1	Change of mean and variances	81
6.4.2	Reduced heterozygosity	83
6.4.3	Inbreeding depression	83
6.5	Management of genetic variation in fish breeding schemes	84

7.	**Selection**	89
	Trygve Gjedrem and Jørn Thodesen	
7.1	Introduction	89
7.2	Natural selection	89
7.3	Artificial selection	90
7.4	Prediction of response to selection	91
7.4.1	Predicting genetic gain, an example	93
7.5	Multiple trait selection	94
7.5.1	Selection for a trait as a ratio	95
7.6	Indirect selection	96
7.6.1	Selection for disease resistance	96
7.6.2	Selection for feed conversion efficiency	97
7.6.3	Selection for increased absorption of nutrient	100
7.6.4	Selection for reduced fat deposition	100
7.7	Effect of selection on genetic variance and covariance	101
7.8	Selection response obtained	102
7.8.1	Disease resistance	102
7.8.2	Growth rate	103
7.8.3	Feed conversion efficiency	104
7.8.4	Date of spawning	105
7.8.5	Sexual maturation	105
7.9	Selection limits	106
7.10	Effect of selection on domestication	108
7.11	Socio-economic effects of improved stocks	110
8.	**Reproductive traits in aquatic animals**	113
	Terje Refstie and Trygve Gjedrem	
8.1	Reproductive cycle	113
8.2	Sex determination	113
8.2.1	Unisexual fishes	113
8.2.2	Hermaphrodites	114
8.2.3	Bisexual fishes	114
8.3	Sex ratio	114
8.4	Sexual characteristics	116
8.5	Age of maturation	116
8.6	Egg and sperm production	117
8.6.1	Production of egg	117
8.6.2	Production of sperm	117
8.7	Fertilization	117
8.8	Genetic determination of sex	118
8.9	Hormone induced sex reversal	118
8.10	Hormone induced ovulation	118
8.11	Preservation of gametes	119
8.12	Influence of reproduction on strategies for genetic improvement	119

9.	**Methods for estimating phenotypic and genetic parameters** Kari Kolstad	121
9.1	Estimating genetic variance components from resemblance between relatives	121
9.1.1	Phenotypic covariances between relatives	121
9.1.2.	Genetic covariance between relatives	122
9.2	Utilising phenotypic covariances between individuals for estimating genetic parameters	123
9.2.1	Estimating additive genetic variance by utilising phenotypic covariance between parent and offspring	124
9.2.2	Estimating additive genetic variance by regression of phenotypic value of offspring on phenotypic value on parents	125
9.2.3	Estimating additive genetic variance by utilising phenotypic covariance between sibs	126
9.3	Mixed model types of analysis	130
9.4	Estimation of heritability	131
9.4.1	Estimating heritability from regression and correlations between relatives	131
9.4.2	Combinating information by REML	133
9.4.3	Heritability of family means and within-family deviations	133
9.5	Estimation of repeatability	134
9.6	Phenotypic and genetic correlations between traits	135
9.6.1	Estimating correlation from covariances between relatives	135
9.7	Correlations estimated from multi-trait models based on mixed linear model techniques and pedigree information	138
9.8	Threshold characters	140
9.8.1	Liability and threshold	140
9.9	Concluding remarks	143
10.	**Breeding strategies** Kjersti Turid Fjalestad	145
10.1	Introduction	145
10.2	Inbreeding	146
10.2.1	Application of inbreeding	147
10.2.2	Inbreeding depression	147
10.2.3	Results from inbreeding experiments	148
10.3	Crossbreeding	150
10.3.1	Heterosis	150
10.3.2	Combining ability	150
10.3.3	Reciprocal recurrent selection	151
10.3.4	Overdominance	152
10.3.5	Diallel cross	153
10.3.6	Three-way and four-way crosses, backcrosses	154
10.3.7	Synthetic populations	154
10.3.8	Application of crossbreeding in fish and shellfish	154
10.3.9	Examples and results from crossbreeding experiments	155

10.4	Pure breeding	157
11.	**Selection methods**	159
	Kjersti Turid Fjalestad	
11.1	Introduction	159
11.2	Individual selection	159
11.3	Pedigree selection	161
11.4	Family selection	161
11.5	Within family selection	164
11.6	Progeny testing	165
11.7	Combined selection	165
11.8	Prediction/Expected response	166
11.9	Correlated response	167
11.10	Indirect selection	168
11.11	Relative merit of the methods	170
12.	**Design of breeding programs**	173
	Bjarne Gjerde	
12.1	Introduction	173
12.2	Base population	174
12.3	Maintenance of additive genetic variance	175
12.4	Inbreeding and risk	176
12.5	Mating design	178
12.5.1	Mass spawning	179
12.5.2	Single pairs	179
12.5.3	Nested	180
12.5.4	Full factorial	181
12.5.5	Partly factorial	181
12.5.6	Future mating design	182
12.5.7	Mating design and inbreeding	182
12.6	Selection	182
12.6.1	Individual selection	182
12.6.2	Within family selection	185
12.6.3	Individual versus BLUP selection	185
12.6.4	Sib selection	187
12.6.5	Progeny testing	188
12.7	Selection and mating	188
12.8	Connectedness	189
12.9	Physical tags versus DNA-markers for parental assignment	190
12.10	Small scale breeding program	191
12.11	Large scale breeding program	192
12.11.1	Breeding nucleus	192
12.11.2	Test animals	193
12.11.3	Selection	193
12.11.4	Multiplier units	193
12.12	Concluding remarks	194

13.	**Prediction of breeding values**	197
	Bjarne Gjerde	
13.1	Introduction	197
13.2	Derivation of predictors	199
13.2.1	Best predictor (BP)	199
13.2.2	Best Linear Predictor (BLP)	200
13.2.3	Best Linear Unbiased Predictor (BLUP)	200
13.3	Selection Index Procedure (SIP)	201
13.3.1	Selection for several traits	201
13.3.2	Selection for one trait	202
13.3.3	Individual selection	203
13.3.4	Within full-sib family selection	204
13.3.5	Sib selection	204
13.3.6	Family selection	205
13.3.7	Own, full and half-sib selection	205
13.3.8	Progeny testing	207
13.4	Accuracy of predicted breeding value	208
13.4.1	Individual selection	209
13.4.2	Within family selection	209
13.4.3	Sib selection	210
13.4.4	Family selection	210
13.4.5	Combined own, full- and half-sib selection	210
13.4.6	Accuracy for different selection methods	211
13.4.7	Prediction error variance	213
13.5	Best Linear Unbiased Predictor (BLUP)	214
13.6	Differences between SIP and BLUP	217
13.7	Prediction of breeding values with BLUP	218
13.7.1	BLUP with A^{-1}, $h^2 = 0.25$	223
13.7.2	BLUP with A^{-1}, $h^2 = 0.50$	225
13.7.3	BLUP without A^{-1}, $h^2 = 0.25$	225
13.7.4	Least squares estimates	225
13.7.5	Breeding values for the progeny	225
13.7.6	The effect of A^{-1}	226
13.7.7	Scaling the predicted values	226
13.7.8	Effects common to full-sibs other than additive genetics	227
14.	**Genotype - environment interaction**	233
	Trygve Gjedrem	
14.1	Introduction	233
14.2	Causes of interaction	233
14.3	Estimation of interaction	235
14.4	Solving the problem of interaction	235
14.5	Estimates of genotype-environment interaction in aquaculture species	235
14.5.1	Carp	235
14.5.2	Salmonides	238
14.5.3	Tilapia	239

14.5.4	Catfish	239
14.5.5	Shrimp	240
14.5.6	Molluscs	241
14.6	Conclusion	242
15.	**Measuring genetic change**	243
	Morten Rye and Trygve Gjedrem	
15.1	Introduction	243
15.2	Control population	243
15.3	Divergent selection	245
15.4	Repeated mating	245
15.5	Use of average breeders	246
15.6	Genetic trend analyses	248
15.7	Prediction of response based on selection paths	248
16.	**Breeding plans**	251
	Trygve Gjedrem	
16.1	Introduction	251
16.2	Fish and shellfish as targets for breeding programs as compared with farm animals	251
16.3	Consequences of level of fecundity	252
16.4	Breeding goal	253
16.4.1	Growth rate	255
16.4.2	Feed conversion efficiency	255
16.4.3	Disease resistance	257
16.4.4	Age at sexual maturation	258
16.4.5	Quality traits	258
16.4.6	Fecundity	259
16.4.7	Recapture frequency	259
16.4.8	Behaviour	259
16.4.9	Summary	260
16.5	Forming a base population	260
16.6	Breeding strategies	261
16.7	Selection methods	262
16.8	Genotype - environment interaction	262
16.9	Identification of animals	263
16.10	Closed life cycle and controlled mating	264
16.11	Mating design	265
16.12	Testing strategies	265
16.13	Registration of records	268
16.14	Systems for selection procedures	270
16.14.1	Individual selection	270
16.14.2	Family selection	271
16.14.3	Progeny testing	272
16.15	Estimation of breeding value	272
16.16	Selection of broodstock	273

16.17 Controlling the amount of genetic gain 273
16.18 Constraints to breeding programs 274
16.18.1 Negative correlated effects 274
16.18.2 Change of breeding goal 274
16.18.3 Spread of diseases 274
16.18.4 Genotype-environment interaction 275
16.18.5 Inbreeding accumulation 275
16.18.6 Reduction in genetic variation 275
16.19 Dissemination of genetic improvement 276
16.20 Combining selective breeding and marker assisted selection 277
16.21 Summary of main elements in a breeding plan 277

17. Organising breeding program 279
Trygve Gjedrem and Terje Refstie
17.1 Practise in farm animals 279
17.2 Organisation of breeding programs in aquaculture species 280
17.2.1 Atlantic salmon in Norway 280
17.3 Structure of breeding programs 282
17.3.1 Fecundity of the species 282
17.3.2 Different breeding goals 283
17.3.3 The chase of sex limited traits 283
17.3.4 Economic value of each animal 283
17.3.5 Wether fertilization takes place internally or externally the female 283
17.3.6 Cost of broodstock production 283
17.4 Farmers cooperatives – breeding companies 284
17.5 Initiatives to start breeding programs 284
17.6 National or local breeding programs? 285

18. Chromosome engineering 287
Terje Refstie and Trygve Gjedrem
18.1 Introduction 287
18.2 Gynogenesis 288
18.2.1 Inactivation of sperm 289
18.2.2 Shock to produce gynogenetic diploids 289
18.2.3 Production of inbred lines 290
18.3 Androgenesis 291
18.4 Triploidy 291
18.4.1 Production of triploids 292
18.4.2 Methods to determine ploidy 294
18.4.3 Sexual maturation of triploids 295
18.4.4 Triploids for fish farming 296
18.5 Production of YY stocks 297
18.6 Tetraploid 298
18.7 Conclusion 298

19. Modern biotechnology and aquaculture 301
Ben Hayes and Øivind Andersen
19.1 Introduction 301
19.2 DNA markers 302
19.3 Linkage maps 303
19.4 Linkage mapping and QTL analysis 306
19.4.1 Detection of quantitative traits loci 306
19.4.2 Strategies to minimise the number of genotypings 310
19.5 Results of QTL detection 312
19.6 Implementing marker assisted selection in breeding programs for aquaculture species 312
19.7 Other applications of DNA markers 314
19.8 Transgenic fish 315
19.9 Future prospects 316

20. Genetic interactions between farmed and wild fish, with examples from the Atlantic salmon case in Norway 319
Hans Bernard Bentsen and Jørn Thodesen
20.1 Introduction 319
20.2 Conservation of biological diversity 319
20.2.1 Diversity of species 319
20.2.2 Genetic diversity 320
20.2.3 Biological diversity 320
20.2.4 Long-term conservation 320
20.3 Introduction of exotic species of fish 321
20.3.1 Effects on the genetic diversity 321
20.3.2 Introduction of micro organisms and parasites 321
20.4 Genetic interactions between genetically improved and wild fish of the same species 322
20.4.1 Genetic differences between farmed and wild populations 322
20.4.2 Gene flow from genetically improved breeds to wild populations 324
20.5 Case study: Interactions between farmed and wild Atlantic salmon in Norway 325
20.5.1 Genetic composition of Norwegian farmed salmon 326
20.5.2 Genetic composition of the wild river populations 327
20.5.3 Gene flow from farmed to wild populations 330
20.5.4 Effects of genetic interactions between farmed and wild Atlantic salmon 332
20.6 Summary 334

Appendix A 335

References 337

Index 361

1. STATUS AND SCOPE OF AQUACULTURE

TRYGVE GJEDREM

1.1 Introduction

About 71 % of the globe's surface is covered by seawater and 1 % percent by fresh water. In 1990 global food production of all types were approximately 4.6 billion metric tons of gross tonnage and about 2.4 billion tons of edible dry matter. Of this 98 percent was produced on land and less than 2 percent came from oceans and inland waters. Plant products constituted 92 percent of human diet and animal products constituted 8 percent diet. Aquaculture production is limited compared to that from land-based animals. In 1966 the production of fish and shellfish was estimated to 1 mill. tons and in 1975 5 million tons (Pillay and Dill, 1979). The production increased to 7.7 mill. tons in 1985 and to 37.9 mill. tons in 2001 (FAO, 2003). Expansion has therefore been rapid in the last decade, Figure 1.1.

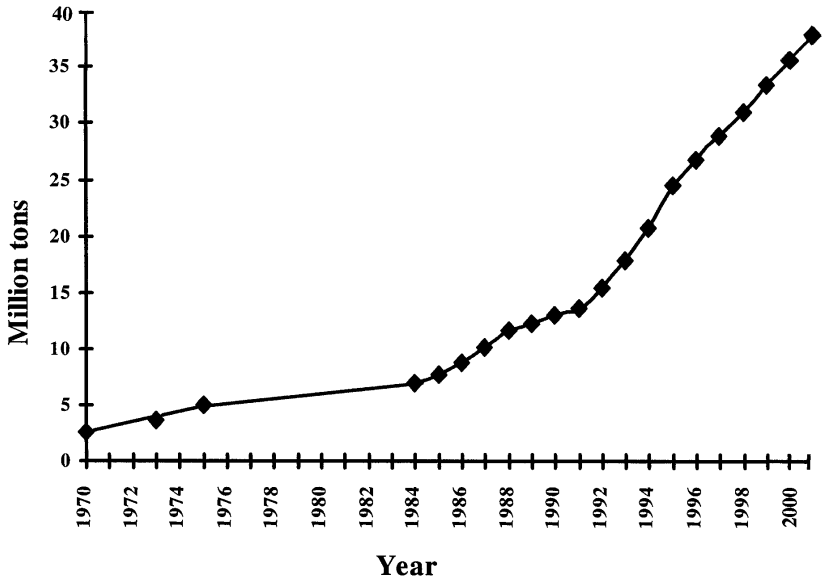


Figure 1.1. World production of fish and shellfish.

The tendency of increased production is expected to continue in coming years (New, 1991). Since capture fisheries is expected to stabilise on 90 - 100 mill. tons, aquaculture production has to increase considerably for the availability of fish and shellfish to be

T. Gjedrem (ed.), Selection and Breeding Programs in Aquaculture, 1–8.
© 2005 *Springer. Printed in the Netherlands.*

maintained. To meat the demand for fish and shellfish, the aquaculture production must, according to New (1991), be 63 mill. tons in year 2025 (Figure 1.2). To reach this goal the production must increase by 4.75 percent per year. The expansion in aquaculture will most likely occur in Asia, especially in tropical and subtropical area where the bulk of the aquatic production takes place today. The increased production is likely to come from more intensive culture, marine farming, sea ranching, and an increased number of species. Limiting factors could be pollution of water particularly freshwater. The availability of fish meal and marine oils could also become a problem.

Efficient breeding programs will be crucial to this development, not only to reach the production goal but also to reduce production cost, improve disease resistance, improve utilisation of feed resources and improving product quality (Gjedrem, 1997b).

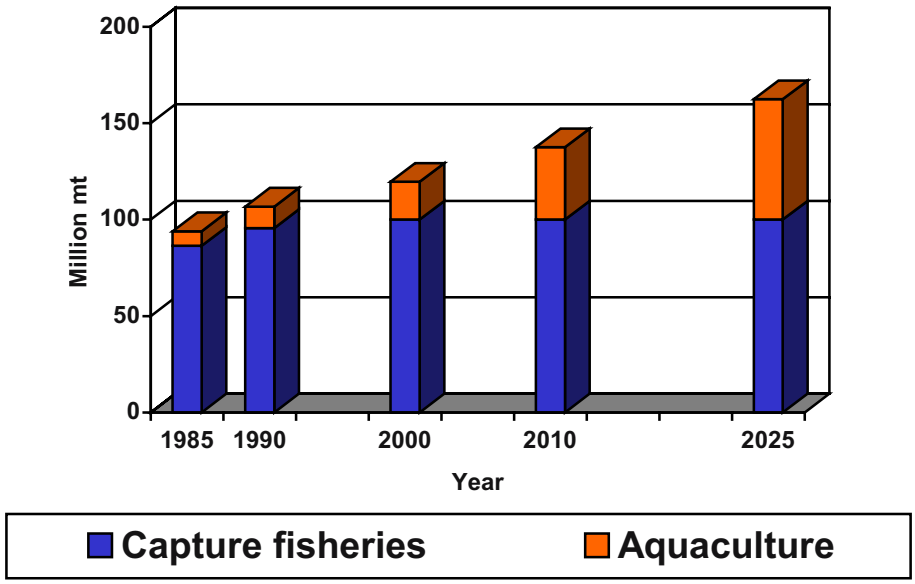


Figure 1.2. Projected requirement for aquaculture production based on population growth and 1989 per capita availability. Reproduced from New (1991) with permission of World Aquaculture Magazine.

1.2 Domestication

Domestication of animals started after the end of the Ice Age. Of the large number of animal species that humans tried to domesticate, very few have been domesticated successfully. Domestication implies several factors and no one factor defines it completely. From Lush (1949) one can conclude that domestication usually means:

- Large changes in behaviour, including tameness, adaptation to captivity, reduced level of stress in the vicinity of humans

- Reproduction and growth is partly under human control
- Animal's products and services are improved
- Domesticated animals depend partly on the farmers care for their existence

Hale (1969) defines domestication as the condition wherein the breeding, care and feeding of animals is more or less controlled by man. It is characterised by genetic changes in behaviour, morphology or physiology, which are brought about by cultivation under artificial conditions. To domesticate a new species is a great challenge. Through natural and artificial selection the animals, which are best adapted to the new, partly human controlled, environment, are selected.

Little is known about how domestication was first accomplished. It is speculated that hunters brought home a few young as pets and that the most gentlest and cooperative were selected and kept. It is uncertain whether some farm animals were descended from single species or from several. Domestication may have taken place in different regions and the domesticated stock may have been crossed or interbred some time later.

In fish domestication seems to have started recently in comparison with land animals. Carp, the oldest domesticated fish, have been farmed for 3-4000 years (Bardach et al., 1972). One theory for the comparatively late domestication of aquaculture species is related to the high fecundity of these species. This theory proposes that because of this high fecundity, only a small number of broodstock were necessary to breed enough offspring for the next generation. As a result, inbreeding depression (and along with it disease and poor performance) accumulated rapidly. Fish farmers were then forced to take new, less inbreed, but undomesticated broodstock from wild, interrupting the cycle of continuous improvement.

An exception is ornamental fish, which are kept in captivity for their aesthetic appeal. The oldest domesticated species is according to Purdom (1993) the goldfish (*Carassius auratus*). These authors believe the goldfish was domesticated in China before 16th century and brought to Japan and later to Europe. Another popular group of ornamental fish is the koi, a colourful derivative of the common carp (*Cyprinus carpio*), and mainly developed in Japan. In this species the large variation in form and colours has been obtained by selection and crossbreeding.

1.3 History of animal breeding

Animal breeding is a body of theory, which aims to address the problem of finding the animals, which produce the best offspring, generation after generation. Variation among animals is essential, and as a consequence we may also say that animal breeding is a theory regarding variation or variance and the utilisation of this variation in breeding and selection.

The history of animal breeding shows that the principles of selection have been applied for thousands of years, long before the theoretical background was developed, with

considerable success. We do not know when artificial selection started to play a role in the domestication of farm animals, but we believe that natural and unintentional artificial selection played a combined role for a long time. Behavioural traits particularly must have been selected for during early stages of domestication.

The modern breeding theory is relatively new, with development beginning early last century. The principles of transmission of genetic material from one generation to the next are the basis of mendelian genetics. The Austrian monk Johan Georg Mendel published in 1866 his classic papers about the laws of particulate inheritance. He developed his theories from experiments with plants. Mendelian genetics deals primarily with qualitative traits controlled by few genes and with discontinuous variation. Mendel's paper was according to Gardner and Snustad (1981) re-discovered simultaneously by three botanists. Hugo de Vries in Holland, is known for his mutation theory and studies on the evening primose in maize, peas, and beans. Eric von Tschermak-Seysenegg in Austria worked with several plants including garden peas. Each of these investigators obtained evidence for Mendel’s principles, from his own independent studies. William Bateson gave this developing science the name genetics in 1905. In 1901 William Bateson provided the first evidence from experiments with chickens that mendelian's “particles” are the basis of inheritance in animals.

The Hardy-Weinberg law from 1908 gave the basic foundation for population genetics, the study of mendelian genetics in populations. It is usually limited to the inheritance of qualitative traits influenced by one or two genes and experimental results are expressed in frequencies of genotypes and phenotypes.

Understanding and development of the theory of quantitative genetics, the inheritance of traits with continuous variation, began at about the same time (quantitative traits are those which are assumed to be regulated by a large number of genes each with small effects). Charles Darwin published his book "The origin of species" in 1859 and focused on the effects of natural selection on accumulation of small changes in traits. The idea of selection to improve quantitative characters is a central conception in quantitative genetics, including the improvement of production traits in livestock and aquaculture species. The biometricians Francis Galton and Karl Pearson contributed greatly to the theory by using the statistical tools of regression and correlation, and discovering the principle that the similarity between an animal and its descendants decreases by one-half from one generation to the next. Later Ronald A. Fisher and Sewall Wright reconciled the apparent discrepancy between the way in which quantitative and mendelian traits are inherited by demonstrating that mendelian frequencies are the basis of biometrical correlations.

Fisher, Wright and J.B.S. Haldane developed their quantitative evaluation of systems of mating and selection procedures and applied their results to problems of evolution, while Wright and Jay L. Lush were the first to make use of these same procedures in the solution of animal breeding problems.

Lush, the father of modern animal breeding, used in the 1930's the theories of Wright and others to establish the methods of estimating the merit of animals for breeding purposes or breeding values. In the early 1950`s following Lush, C.R. Henderson and Alan Robertson developed methods and programs for the estimation of breeding values of dairy sires using computers.

Today, research institutions in most countries are continuously working on problems related to selective breeding in order to develop new and more efficient methods for improving existing breeding practices and, consequently, increase the rate of genetic improvement of the important livestock species in their respective countries.

1.4 Aim of breeding - Why change a population?

The main purpose of a breeding program is to develop animals, which more efficiently utilise available feed, land and water resources. It is very probable that by pursing this goal we will also select animals, which are better adapted to the captive environment, as they must produce well under these conditions. It is easy to observe that undomesticated animals, including fish and shellfish, are continuously stressed in captivity. Therefore there is also an animal welfare dimension to the question of whether to domesticate a species or continuously take representatives from the wild. In fact the aim of a breeding program is two fold:

- To increase the welfare of farmed animals
- To increase productivity of animals and to improve their product quality

1.5 Effect of breeding programs in farm animals

1.5.1 WHAT HAS BEEN ACCOMPLISHED?

It is not known how fast farm animals were domesticated. However, it is evident today that there are large differences between domestic animals and their wild relatives concerning behaviour, conformation and production traits. During the process of domestication the animals became more and more adapted to captivity and it is likely that the most calm and productive animals were preferred as breeders. To what extent the domesticated animals have changed compared to the wild and what has been the genetic change is not known.

The first efficient breeding programs for farm animals were developed during the 1930's, following the development of the theoretical base for quantitative genetics, which was not known before. Since then the improvement of production traits have been remarkable. As illustrated in Figure 1.3 the average number of eggs laid per year by a hen steadily increased from approximately 120 in the 1940s to more than 320 by the mid-1980s; the average milk production per cow in a single lactation of 305 days increased from approximately 2 000 kg in 1945 to more than 5 000 kg by 1980. Average daily gain in the pig industry has increased from about 450 g around 1960 to 800 g in

the 1980s. The production performance in farm animals continues to increase and there is no sign that the genetic gain levels off. According to Barlow (1983) and Michell et al. (1982) the estimated cost/benefit ratios of genetic programs in farm animals is somewhere between 1:5 to 1:50, depending on the species.

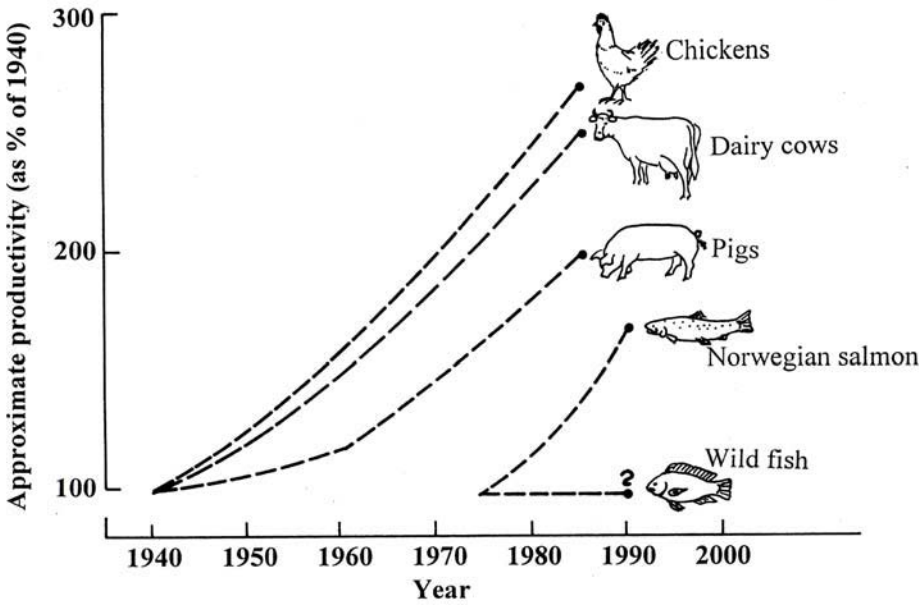


Figure 1.3. Genetic research initiated in the 1940's has resulted in remarkable developments in increasing the productivity of domestic mammals and birds. Reproduced from Eknath et al. (1991) by permission of World Fish Center.

This increase in productivity has led to a much better utilisation of available feed resources and cost of production has been dramatically reduced. It would be unimaginable today to use wild stocks for the production of animal products such as eggs, meat, pork, milk and wool. The situation in many aquaculture species, however, is quite different.

1.6 Breeding in aquaculture

In comparison with livestock species, application of quantitative genetic principals to fish breeding has been limited until very recently. Fish farming started 4000 to 5000 years ago in China. Artificial hatching of fish was already practised in China around year 2000 BC (Lin, 1940). The earliest written descriptions of pond culture are from China during the 12th century BC (Menasveta and Fast, 1998). The development of fish culture grew parallel with the culture of silk worms, as the pupa of the silkworm and their faeces provided supplementary feed for fish (Hickling, 1962). The oldest book on

aquaculture is from China by Fan Lee and was published as early as in the year 475 B.C (Ling, 1977). Fan Lee described the aquaculture techniques practised around 500 BC. It is not known when domestication of fish started in China. One can still meet carp breeders in China saying, «To get strong breeders we sample wild broodstock from the big rivers», indicating that artificial selection to improve production traits does not currently take place to a large degree. According to Schaperclaus (1961) common carp farmers in Europe have been selecting the largest fish for hundreds of years compared with thousands of years in livestock species. Some experiments to improve disease resistance and growth rate started in the 1920's (Embody and Hyford, 1925; Lewis, 1944; Donaldson and Olson, 1955).

Although fish farming is old in several parts of the world it is not obvious to what degree domestication has taken part. In Europe and Israel common carp have been bred in closed populations for a long time (Schaperclaus, 1961). Rainbow trout *(Oncorhynchus mykiss)* have been kept in captivity for many generations in Europe since the 1890's and show clear sign of tameness compared to the wild type. Tilapia were first found in Indonesia in 1937 and farming started after the second world war in South East Asia.

In Norway farming of Atlantic salmon *(Salmo salar)* started in the late 1960's and the majority of the farmed fish today has been artificial selected for 7 generations and show many signs of domestication. This effect is seen most markedly in the freshwater phase as reduced cannibalism, wildness of fry and fingerlings, which means that the behaviour of the fish has changed and they use less energy to swim, fight and try to hide behind each other.

Milkfish *(Chanos chanos)* and other species, including marine shrimp were commonly cultured on a large scale in Indonesian brackish water ponds by the 15th century AD. However, important aquaculture species like milkfish, yellowtail *(Seriola quingueradis)*, crustacean and molluscs rely largely to date (2004) on wild broodstock or fry.

1.7 Present status of breeding in aquaculture

In aquaculture, selection programs are not commonly used by the industry and for many species production still rely completely on catching wild broodstock and/or fry.

There is no obvious reason for the lack of efficient breeding programs in aquaculture. The economic important traits in fish and shellfish appear to be little different from those in farm animals and plants. Selection response is usually higher in fish and shellfish than in farm animals (Olesen et al., 2003).

One reason for the scarcity of breeding programs in aquaculture species is that the reproductive cycle is often complex, and so is frequently not fully understood, and is therefore not able to be completed or controlled in captivity. This is the case particularly for marine species. A further factor contributing to the scarcity of breeding programs in aquaculture species may be the deterioration of the stock simply because of a rapid build

up of inbreeding as a result of using few broodstock each generation (and without identification to prevent re-use). This is a problem in all species with high fecundity. Other reasons may be that researchers, extension personnel and fish farmers are not well educated in breeding theory, since education in fish biology involves little or no attention to quantitative genetics and breeding plans. The majority of published papers discussing genetics in aquaculture deal with genetic variation in qualitative traits and population genetics.

Because there has been little interest in developing breeding programs in aquaculture, the information about phenotypic and genetic parameters of economically important traits are quite limited for most of the species farmed. Before a breeding program can be established, breeding goal must be defined, estimates of genetic variance, heritability, phenotypic and genetic correlations among traits must be available. Some of the first estimates of heritability in aquatic species were published by Aulstad et al. (1972) for body weight and body length in rainbow trout fingerlings and by Kirpichnikov (1972) for body weight of common carp fingerlings. Later a large number of reliable estimates of genetic and phenotypic parameters were published particularly for salmonid fishes. However, data is still lacking for a number of important species. There is therefore a great need to run breeding experiments in order to get reliable estimates of genetic parameters for economically important traits in the most important farmed species.

The following chapters describe the basic principles of selective breeding programs with the aim of leading to an understanding of how knowledge of mendelian, population and quantitative genetics can be used for the permanent improvement of fish and shellfish.

2 BASIC GENETICS

TRYGVE GJEDREM AND ØIVIND ANDERSEN

2. 1Molecular genetics

2.1.1 DNA - MOLECULE

The smallest unit in an organism is the cell. The cell has its own metabolism, it can grow, multiply and respond to external influences. The cell is surrounded by a membrane, which makes it possible for substances to pass in and out. The nucleus of the cell is surrounded by the cytoplasm, which is a semi fluid mass and consists of organic material, particularly protein. In addition, the cytoplasm contains several particles called organelles: mitochondria, Golgi apparatus, ribosomes, lysosymes and so on. In the

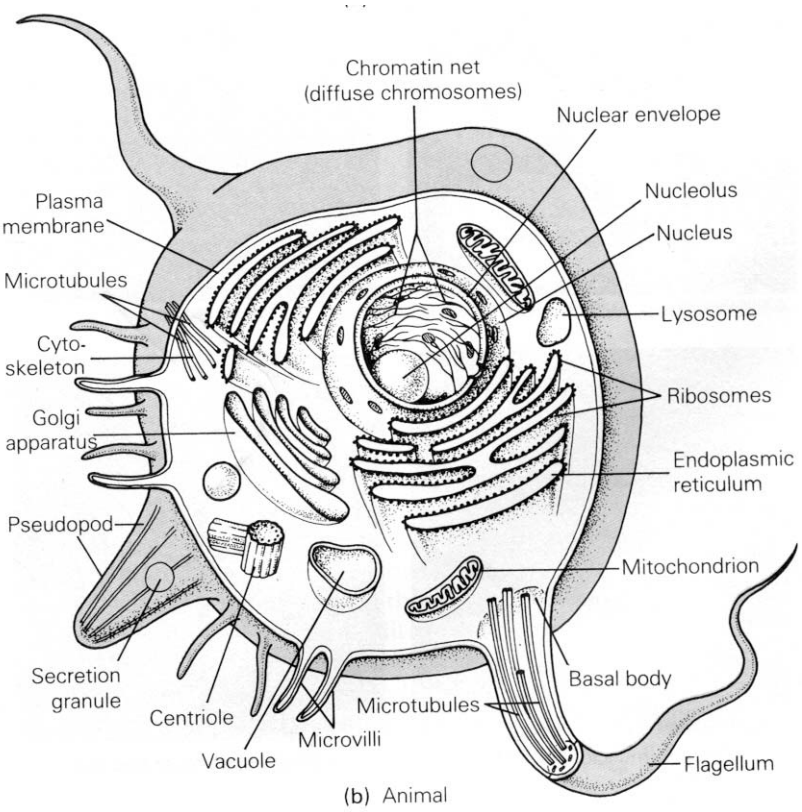


Figure 2.1. A cell with some selected structures. Reproduced from Wallace et al. (1986) by permission of Pearson Education Inc.

T. Gjedrem (ed.), Selection and Breeding Programs in Aquaculture, 9–21.
© 2005 *Springer. Printed in the Netherlands.*

mitochondria there is some genetic information, but its effect is directed by the nucleus (Figure 2.1).

The nucleus is of particular interest since it contains all the genetic information about the development, metabolism and behaviour of the organism. Every cell of an individual has this complete set of genetic information stored in the DNA-molecule. The DNA is organized in threadlike bodies called chromosomes within the nucleus membrane. The chromosomes are rich in nucleoproteins, which are basic proteins that bind to nucleic acids. Nucleic acids are of two types depending of the attached sugar molecules: Deoxyribose nucleic acid (DNA) and ribose nucleic acid (RNA) (Figure 2.2). For some time it was not obvious which of these was the carrier of the genetic information, until it was established that DNA is the basic substance for inheritance. In some simple viruss without DNA, RNA is the genetic material.

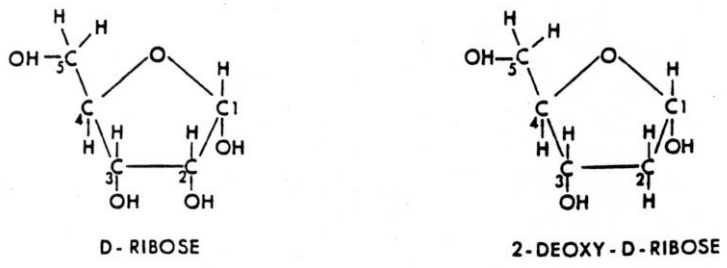


Figure 2.2. Chemical structure of D-Ribose (RNA) and 2-deoxy-D-ribose (DNA). The five carbons are numbered 1-5.

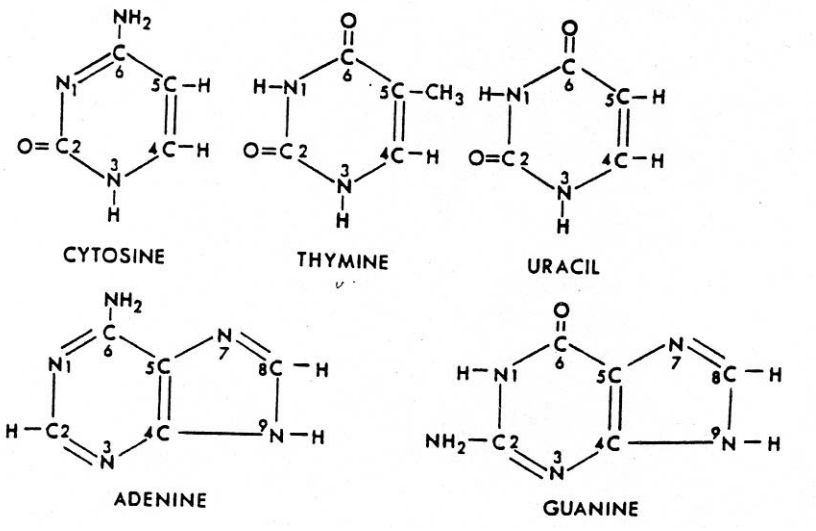


Figure 2.3. Chemical structure of the bases included in the DNA and RNA.

Watson and Crick (1953) showed that the structure of DNA molecule was like a ladder twisted into a right-handed double helix. The uprights of the ladder consist of alternating phosphate (P) and deoxyribose sugar (S) groups. The cross rungs consist of paired combinations of four different base: Adenine *(A)*, guanine (G), thymine (T) and cytosine (C), Figure 2.3. Due to hydrogen bindings each base pair always consists of a purine (A or G) and a pyrimidine (C or T) in four different combinations:

C + G G + C A + T T + A

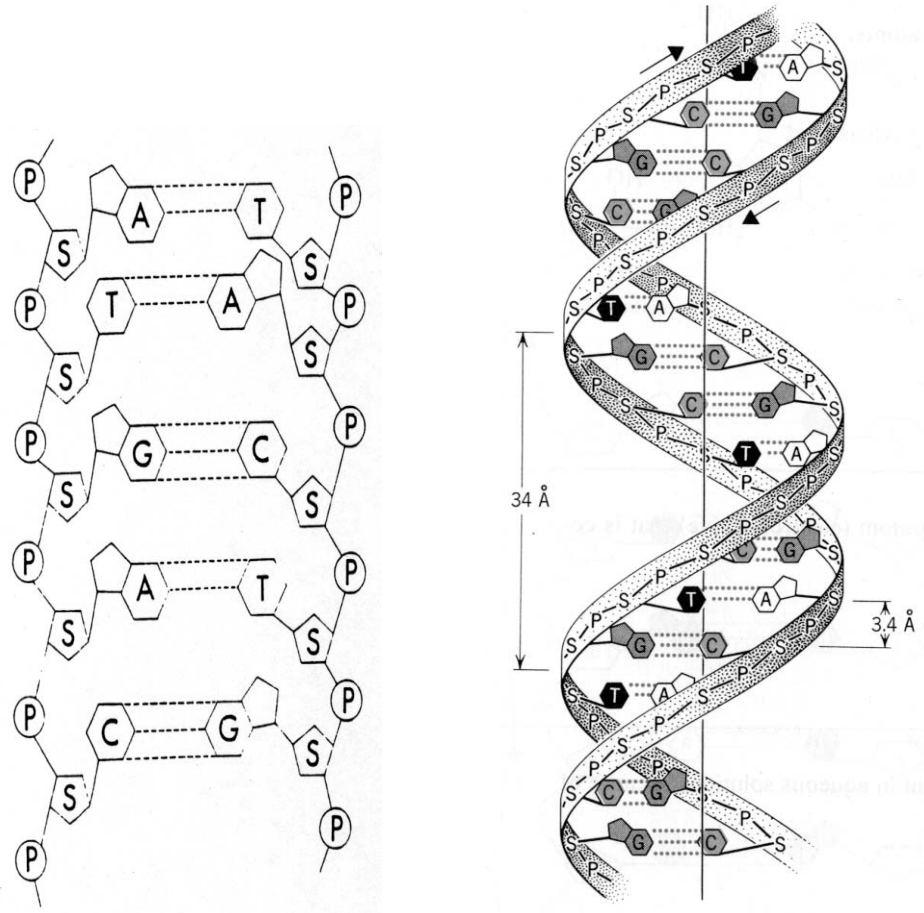


Figure 2.4. Left: Base pairing between the two DNA strands. Right: The double helix structure of DNA. Reproduced from Gardner and Snustad (1981) by permission of John Wiley & Sons, Inc.

RNA differs from DNA in having the nucleotide base uracil (U) instead of thymine. The DNA molecule has 10 base pairs in each full turn of the helix. The distance between each pair is 3.4 Å (1 Å= ten million part of a mm), Figure 2.4. If a DNA molecule in an

animal cell was stretched out, it would be more than two meters long. With different sequences of base pairs the genetic information, which can be stored is almost infinite. It has been calculated that a virus has 5000 base pairs in the DNA molecule. The information (sequence) contained in this length of DNA could be written on one page of a book. A cell of bacteria contains 1000 times more information or 1000 pages in a book. In a mammalian cell the genetic information is so large that 1 million pages would be needed to write the information down. The DNA molecule is organised into structures called chromosomes, which appear in pairs, one chromosome inherited from the male parent and one from the female parent. The number of chromosomes varies among species and in fact chromosome number can be used to characterise species, Table 2.1. A gene is a small section of DNA sequence, which usually codes for the formation of a protein. The length of sequence representing a gene may vary from a few hundred to several thousand base pairs. Although each cell of an organism contains all the chromosomes and all genes, some types of cells utilize different genes at different stages of development and in response to different stimuli. This results in cell differention in different tissues. The estimated number of genes varies between animals, but for most vertebrates is probably it in the order of 28 000 – 34 000 (Crollius et al., 2000). In addition to the recipe for a protein, genes also consist of a promotor, which is found in front of the protein recipe and is to begin protein synthesis. The promotor is also responsible for determining in which cells the gene should function, at what time and in what quantity. The genes are interspersed by inactive DNA segments, which are believed to make up as much as 85 - 90 % of the DNA molecule.

Table 2.1. Number of chromosomes in some aquaculture species.

Species	No. chromosomes	Species	No. chromosoms
Atlantic salmon	58	Pink salmon	52
Brown trout	80	Chum salmon	74
Sea trout	80	Sockeye salmon	56-58
Arctic char	80	Chinook salmon	68
Brook trout	84	Coho	60
Common carp	98-104	Rainbow trout	60
Gras carp	48	Tilapia *(O. niloticus)*	44
Rohu carp	50	Tilapia *(O. mosambicus)*	44
Mrigal	50	Channel catfish	58
Catla carp	50	Turbot	88
Silver carp	48	Striped snake head	40
Bighead carp	48	Shrimp *(Penaeus)*	86-92
Crusian carp	94-100	Prawn *(M. rosenbergii)*	118
Mosquito fish	48	Mussels	28

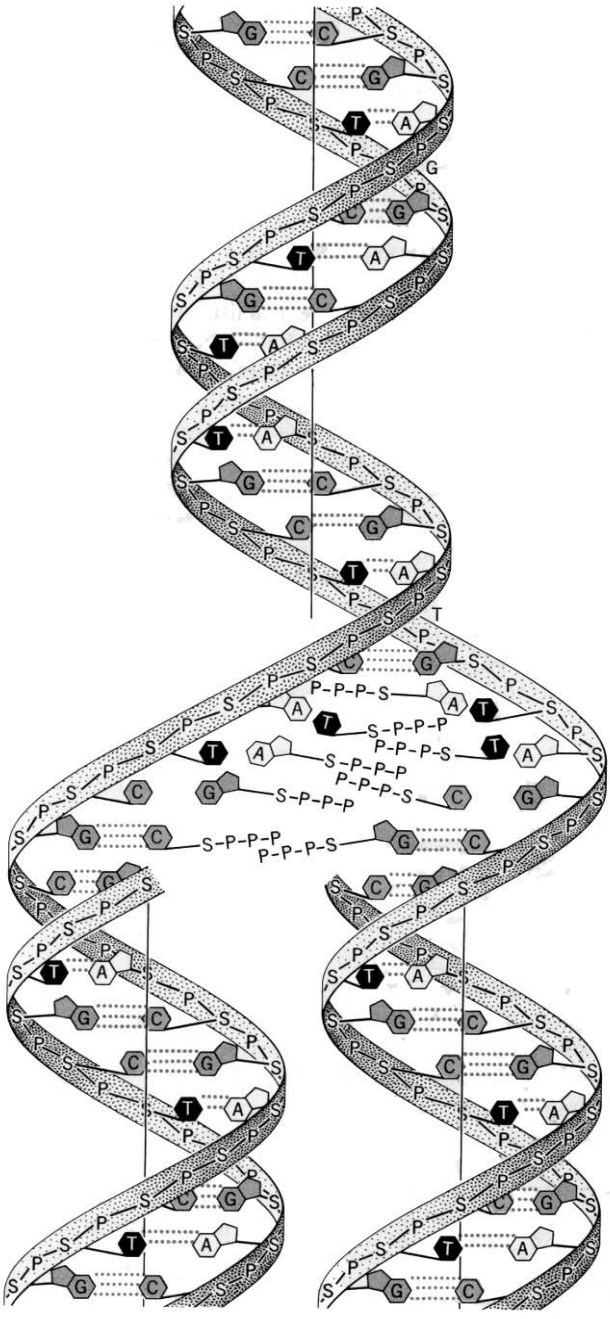


Figure 2.5. The model proposed by Watson and Crick for the replication of the DNA double helix. Reproduced from Gardner and Snustad (1981) by permission of John Wiley & Sons Inc.

2.1.2 REPLICATION OF DNA

Before a cell divides into two identical cells, the DNA-molecule must produce a copy itself. Watson and Crick (1953) uncovered the mechanism of DNA replication. A number of enzymes are active in the replication process. One enzyme begins the process by breaking the bond keeping the base pairs of the double helix together. The strands are divided in the middle and the two helixes are separated as illustrated in Figure 2.5. At the same time as the division process is occurring, the nucleotides of the helix are replicated. These nucleotides will form a complementary helix so that each daughter DNA-molecule will consist of one intact strand from the parental double-stranded helix and one newly synthesised complementary strand. Hence, DNA replication is a semi conservative process.

There two mutation processes, which can result in non-perfect replication. These processes are base-pair substitutions and frame shift mutations. Less than 20 % of spontaneous mutations are believed to be due to base-pair substitutions. Most of the rest are frame shift mutations in which one or several nucleotides are inserted or deleted. The most serious effect of error in copying of DNA molecule is when it occurs in production of egg and sperm because it may be transmitted to the next generation.

2.1.3 PROTEIN SYNTHESIS

Proteins play a central role in all living organisms. Variation in life and life processes is in general a reflection of variation and complexity in proteins. Every protein has a special function, which depends on the protein structure. The conformation of a protein is determined by the combination of the 20 possible amino acids, which it contains. The DNA sequence is not translated directly into amino acids and subsequent protein. In fact the DNA molecule is translated into a RNA molecule before the protein synthesis begins. This RNA molecule then passes the nucleus membrane and reaches the area in cytoplasm where the protein synthesis takes place. Thus it is this messenger RNA (mRNA), which programs the formation of the protein. The flow of information in the protein synthesis follows this path:

$$\text{DNA} \rightarrow \text{RNA} \rightarrow \text{Protein.}$$

Enzymes, which are themselves proteins, are important catalysers in the protein synthesis. These enzymes are very efficient, specialised and direct only one or a few chemical processes.

2.2 Cell division

Cell division is an endless repetition of growth, chromosomal replication, mitosis and cytoplasmic division. The cell cycle is traditionally divided into four phases, mitosis and three phases called interphases. In the first interphase the chromosomes decondense as each daughter cell of the previous division enter this phase. This phase is a very active period where proteins are synthesised, and the chromosomes consist of only a single

DNA strand. In the second phase, which lasts for some hours, the DNA is replicated. In the third phase the cell prepares for mitosis. The proteins of the mitotic spindle, which is involved in the movements of chromosomes during mitosis, are synthesised. At this stage the chromosome consists of two DNA molecules and their related proteins joined at the centromeres are called a chromatid.

2.2.1 MITOSIS

Mitosis is divided into four sub phases (eg. Wallace et al. 1986):

1. In prophase, the chromosomes condense, the nucleoli and nuclear membrane fade from view and the mitotic spindle apparatus forms
2. In metaphase, the chromosomes are brought to a well-defined plane in the middle of the mitotic spindle, which is attached to the chromosome at the centromere
3. In anaphase, the centromere of each chromosome separates and the two daughter chromosomes of each pair travel to the opposite ends (poles) of the spindle
4. In telophase, new nuclear membrane forms around each group of daughter chromosomes, the nucleoli appear, the chromosomes decondense, and finally the cytoplasm of the cell divides to form two daughter cells, Figure 2.6

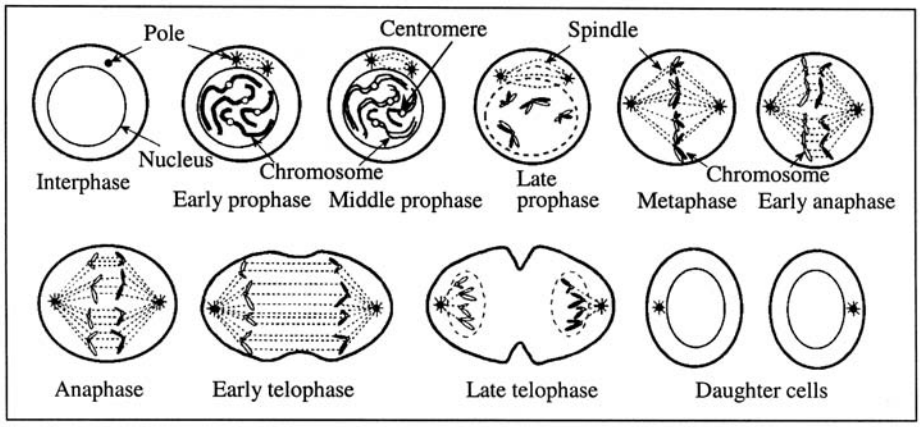


Figure 2.6 Cell division, mitosis. Reproduced from Gjedrem (1995a) by permission of JUUL Forlag.

2.2.2 MEIOSIS

The animal cell is diploid with two sets of chromosomes and genes, originally one set from the dam (mother) and one set from the sire (father). During the process of meiosis diploid stem cells give rise to the haploid gametes (eggs and sperm), but as the gametes join together at fertilisation the progeny will be diploid.

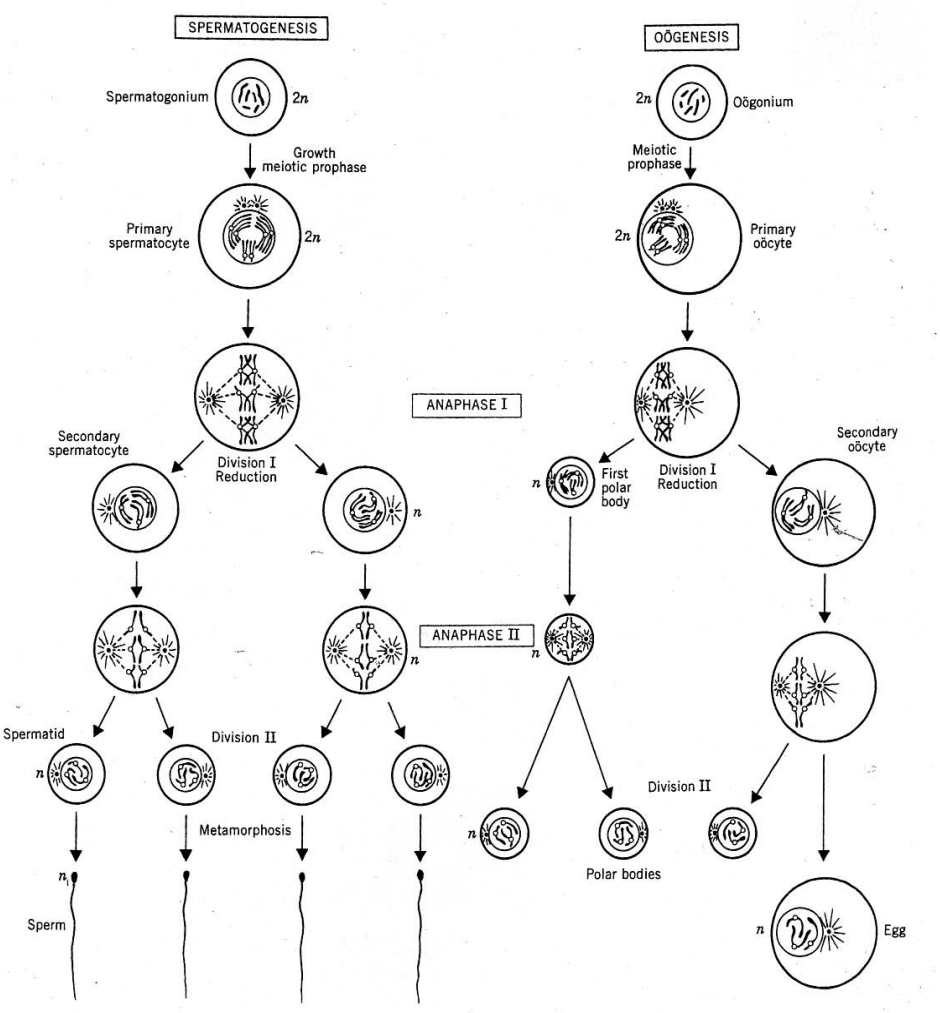


Figure 2.7. Cell division, meiosis. Spermatogenese right and oogenese left. Reproduced from Gardner and Snustad (1981) by permission of John Wiley & Sons Inc.

Meiosis consists of two cell divisions with two prophases, two metaphases, two anaphases and two telophases. Before meiosis begins, the chromosomes are replicated into two sister chromatides. Each pair of sister chromatides remains together throughout the first meiotic division. No further DNA replication occurs in meiosis. In prophases I, the homologous chromosomes come to lie alongside each other, so that each member of the pair consists of two sister chromatides. In metaphase I, synaptic pairs of chromosomes line up at the equatorial plane of the cell. In anaphase I, the member of each pair move towards opposite poles so that each half of the cell contains one chromosome of each type. Thus, the nuclei that form in telophase I are genetically

haploid, but contain two copies of each sister chromatid. In the following interphase period there is no DNA replication.

The second meiotic division for sperm is similar to mitosis. The sister chromatides separate at their centromeres and move to opposite poles and four haploid sperms are produced.

Meiosis in females results in only one egg. The other cells are lost as polar bodies following unequal cleavages in telophase I and II. In telophase I cytokinesis (division of cytoplasm) result in producing a large cell and a tiny functionless polar body. In fish the second meiotic division will not be completed before fertilization has occurred. Sperm and egg will fuse and a second polar body will be excluded, Figure 2.7.

2.2.3 LOCATION OF GENES

The location on a chromosome is called a locus (plural, loci). In a diploid animal there are homologos pairs of chromosomes, each chromosome in the pair containing a complete set of genes (apart from in the case of the sex chromosomes in some animals). The alternative forms of genes or non-coding sequence at a locus are known as the alleles of the locus. There may be a great deal of genetic variation among normal alleles. Often multiple alleles can occur at a given locus although any one animal can have only two alleles - one from each parent.

2.3 Mendelian inheritance

The Austrian monk Johann Gregor Mendel (1822-1884) discovered the basic rules of heredity. The background was a series of well planed crossbreeding experiments and mathematical calculations of the result, which he published in 1866. Mendels two laws of inheritance are:

First law: Inheritance is by units or particles, now called genes, which maintain their identity from generation to generation. It is also known as the law of segregation of genes.

Second law: These units are present in duplicate in each individual. It is also called the law of independent assortment, which means that segregation at one locus does not influence segregation at another.

Mendels statement that the inheritance is units or particles is fundamental to genetics. He thought that all genes segregated independently. Later it was shown that because of linkage not all genes segregate independently. He discovered dominance and commented on it.

When an individual reproduce it transmits to each offspring one or the other of the genes in each pair it possesses. Thus the parents give to its progeny only a sample of half of its own genes. The laws of chance govern this sampling. Mendel’s findings showed that Darwins theory of blending inheritance was wrong. Darwin believed that the «blood» or

traits that are inherited from both parents, blended in the offspring, just as two colours of ink blend when they are mixed.

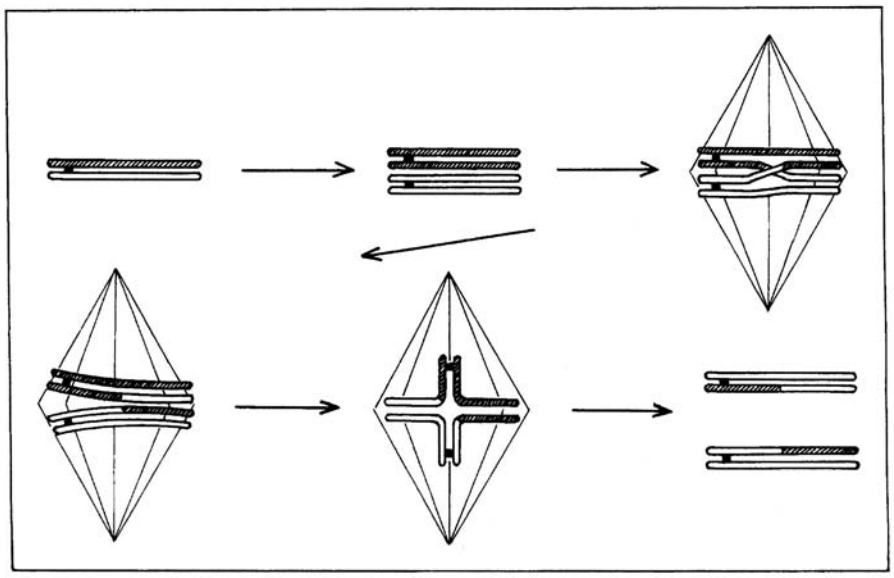


Figure 2.8 Crossing over. Reproduced from Gjedrem (1995a) by permission of JUUL Forlag.

Genes on the same chromosome are expected to be linked in inheritance. There is, however, a process by which homologous chromosomes can break at corresponding sites and exchange parts, thus leading to recombination of the genes located on these chromosomes. Such an exchange is called crossover and is illustrated in Figure 2.8. The result of this meiotic event is that some of the offspring receive recombinant gametes from their parents. As the length of chromosome between two genes increases, the more likely it is that there will be a recombination between them. The frequency of crossovers is therefore a function of the distance between the genes. In practise, the number of crossover events in a population of chromosomes, between two genes, can be used to estimate the distance between the genes. Double crossovers can also occur, though such events are rare between closely linked genes.

2.3.1 EXAMPLE OF MENDELIAN TRAITS IN FISH

2.3.1.1 The inheritance of colour

There are examples of traits in fish, which follow simple Mendelian inheritance. The occurrence of albinism is a good example. Albinism has been observed in a number of species: In channel catfish by Bondari (1984); in rainbow trout by Bridges and von Limbach (1972); in medaka by Yamamoto (1969a). These investigations have found that albinism is caused by a single recessive gene *(a)* and that presence of the dominant gene (*A*) causes a coloured body. This is illustrated in Figure 2.9. A true albino (*aa*) rainbow trout has yellow body colour and the eyes are without melanin (appearing pink rather than the normal black). It should be mentioned that the flesh from the albino fish can be coloured reddish by feeding the fish with astaxanthine.

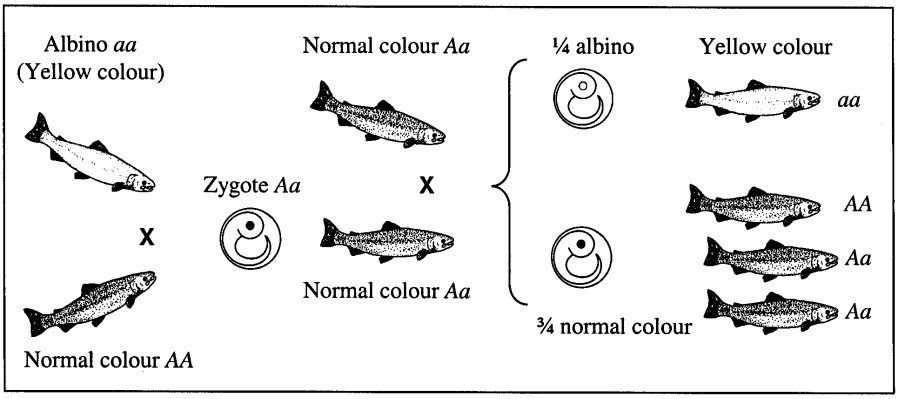


Figure 2.9 Inheritance of albinism in rainbow trout.

2.3.1.2 Inheritance of scale pattern
In common carp there is large variation in the scale pattern, which is controlled by genes in two loci. The S locus gene controls the scaliness and the N locus gene modifies the pattern to show a linear scale pattern. Thus there is an epistatic interaction between these two loci. There are basically four scale pattern phenotypes in common carp:

1. Scaly carp also called wild carp
2. Mirror carp with some scales scattered around the body
3. Linear carp with scales in linear arrays along the body
4. Leather carp with very few scales.

All these forms are found both in wild as well as in farmed common carp. The genotypes for these groups are as illustrated in Figure 2.10:

SS nn or *Ss nn* - scaly carp
ss nn - scattered or mirror carp
SS Nn or *Ss Nn* - linear carp
ss Nn - leather carp
*SS NN*or *Ss NN* or *ss NN*- not viable

By crossing linear and leather carp there will be a mortality of 25 %, which is a problem and should be taken into account in carp farming. According to Kirpichnikov (1981) the growth of scaled carp is generally greater than that of the scattered individuals, particularly if the fishes do not obtain additional food. Linear and leather carp grow more slowly than the others.

2.4 Gene action

Each active gene has a certain effect. Today much research is directed towards studying single gene effects, and localizing or mapping the genes to the chromosomes. This type of work will contribute to our knowledge of how genes function. We do know that the function of genes is strictly regulated and they are switched up and down in different cells and at different times. Development of a zygote into an embryo and finally to a mature individual is a precisely organised process. The cells differentiate into highly specialised tissues and organs, which maintain their diverse functions throughout the life of the individual. Thus the need for specific genes to be active varies during the life cycle.

The simplest form of gene effect is that the value of each gene is additive and that the sum effect of two genes in a locus is the average of the two. For quantitative traits of economic importance there are in general many loci behind each trait and the majority of the genes have small effects.

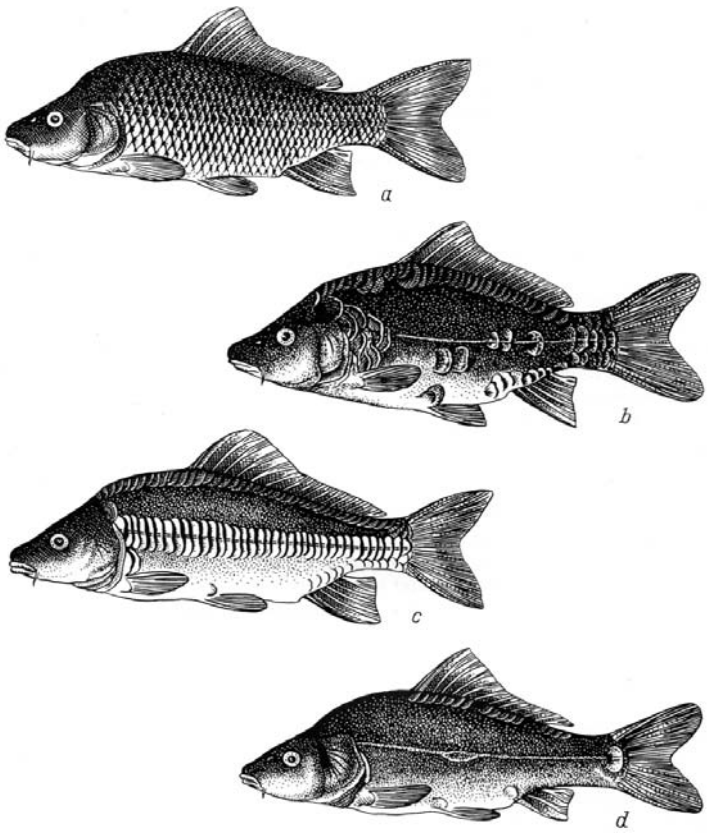


Figure 2.10. Types of scaling in the common carp. **a** scaled (*SSnn* and *Ssnn*); **b** scattered or mirror (*ssnn*); **c** linear (*SSNn* and *SsNn*); **d** leather (*ssNn*). Reproduced from Kirpichnikov (1981) by permission of Springer-Verlag GmbH.

In many cases the effect of one allele of a gene is dependent on the effect of the other allele at the gene of the same locus, termed a dominance interaction. Let us take a look at the heterozygote *Aa.* There are at least three possibilities for value of this genotype as illustrated in Figure 10.2. When genotype *Aa* is equal to *AA* genotype, *A* is dominant. This interaction occurs frequently in qualitative traits, such as albino colour in rainbow trout. When the heterozygote is superior to the homozygotes (either *AA* and *aa*) the interaction is called over-dominance. When the heterozygote is average of the

homozygotes there is no interaction and the inheritance is termed intermediary or additive.

Interaction between the alleles of genes at different loci is termed epistasis. This is due to that one gene in one locus influences the effect of a gene in another locus. In poultry, gene I has epistasitic control over the colour gene *C*. Birds with the following genotypes are white: *IC, Ic* and *ic* while *iC* are coloured. Birds with the dominant gene *I* is white even if they have the dominant gene *C* and birds may be white whether they are recessive in *cc* or dominant in *I* (Crawford, 1990).

The large number of genes in vertebrates, and the possibility of polymorphism at these genes, makes it evident that an enormous number of different kinds of individuals are possible. From Lush (1949) we quote « A comparison show the enormous number of different kinds of individuals possible is furnished by the physicists' estimate that the number of electrons in the universe is about 10^{80}. In a species in which only 200 pairs of genes are heterozygous there could be 10^{95} different kinds of individuals. This is a million billion times as many as there are electrons in the universe». In this example only two alleles were considered, however, if m alleles are present in each of n allelic series the number of different kinds of gametes possible is m^n and the number of different kinds of genotypes possible is $[m(m+1)/2]^2$ (Lush, 1949).

3. POPULATION GENETICS

ØIVIND ANDERSEN AND BEN HAYES

3.1 Introduction

The aim of fish breeding is not to change individual fish, but rather the fish population. Thus, our knowledge about Mendelian inheritance must be extended from the level of the individual to the level of the population. Population genetics may be defined as the study of Mendelian genetics in a group of interbreeding individuals who share a common gene pool. The number of genes in this pool depends on the species, ranging from about 4000 in the bacterium *E. coli* to about 30.000 in the vertebrate genome. Due to the accumulation of mutations over very long periods of time, the DNA sequence of these genes can vary slightly among individuals in the population. These different forms of the genes are called alleles. If the different alleles result in differences of the gene products, the result can be genetic variation for that trait within the population. The proportions of the different alleles for each gene in the pool, the allelic frequencies, determines the genetic characteristics of the population. The product of expression of these alleles interacts with the environment in which the population exists. Some of these interactions give the individuals carrying a particular allelic conformation an advantage over other combinations. Natural selection results in an accumulation of these favorable alleles in the population and leads to a change in allele frequency.

3.2 Hardy-Weinberg equilibrium

In 1908, the English mathematician G. H. Hardy and the German physician W. Weinberg independently discovered that in a large interbreeding population, allele frequencies will remain constant from one generation to the next (they assumed that mating was at random, and there was no selection pressure or other factors altering allele frequencies). Such a population is said to be in Hardy-Weinberg equilibrium.

To understand how Hardy-Weinberg equilibrium occurs, consider a diploid organism, in which the somatic cells have two copies of each gene located in a specific position, or locus, on the pair of homologue chromosomes received from mother and father. Mendels first law states that in a sexually reproducing organism the two alleles separate, or segregate, during the production of gametes. Genotype frequencies may then be calculated from the allele frequencies because the proportion of an allele in a population equals the proportion of gametes containing this allele. Thus, the genotype frequencies are entirely determined by the allele frequencies in the population. This is formulated in the Hardy-Weinberg law stating that the genotype frequencies are maintained constant in a population as long as the allele frequencies do not change and so long as mating occurs at random.

The example of skin color in tilapia demonstrates the consequences of Hardy-Weinberg equilibrium. Pigmentation in tilapia is under control of a single gene, existing in only

T. Gjedrem (ed.), Selection and Breeding Programs in Aquaculture, 23–33.
© 2005 *Springer. Printed in the Netherlands.*

two allelic forms *A* and *a.* Let the frequency of *A* in the population be p and the frequency of *a* in the population be q, and p+q=1 (as there are only two alleles in the population). The normally pigmented black tilapia is homozygous *AA*, the heterozygous *Aa* is bronze pigmented, and the gold individuals are homozygous *aa* (Figure 3.1). If the frequency of the allele *A* is set to 0.7 in a tilapia population, the probability of any sperm or egg (gamete) bearing an *A* is 0.7, as a consequence of Mendel's first law. Then, the probability of bringing two *A*-bearing gametes together at fertilization is p x p = p^2 = 0.49. So the frequency of homozygous *AA*, or black individuals, is 49 % in the population. Likewise, since the frequency of the alternative allele *a* is 1 - 0.7 = 0.3, the probability of a gamete bearing *a* is 0.3. Thus, the probability of joining two *a*-bearing gametes is q x q = q^2 = 0.09. So 9 % of the fish in the population are homozygous *aa* (these fish are gold pigmented). The frequency of the heterozygous individuals (gold pigmented) is pq + qp = 2pq = 0.42, or 42%. Altogether, the frequencies of black, bronze and gold individuals add up to all of the individuals in the tilapia population, and $p^2 + 2pq + q^2 = 1$.

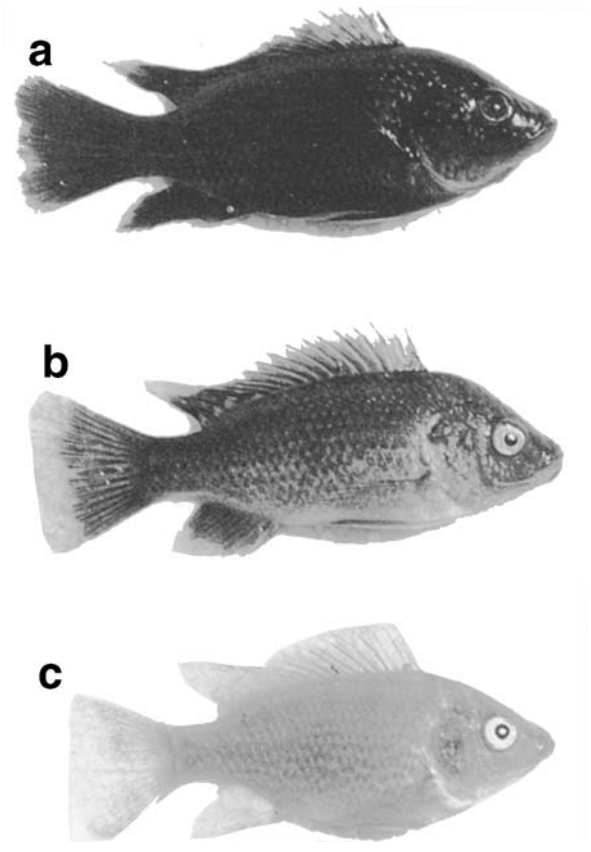


Figure 3.1. Pigmentation in tilapia. Homozygous *AA* is black (a), heterozygous *Aa* is bronze (b) and homozygous *aa* is gold (c). Reproduced from Tave et al. (1989) by permission of J. World Aquaculture Society.

The picture becomes more complicated when alleles of multiple loci are considered. In the case of two allele pairs *A/a* and *B/b* of two loci located on the same chromosome, they will not segregate independently upon gamete formation. Hence, for any population not already in equilibrium, linkage between the loci will delay the attainment of equilibrium. This is referred to as linkage disequilibria. Many generations of random mating may be required before equilibrium is attained, with more generations of mating required for more closely linked loci (as the chance of recombination between the two loci is lower).

Tests for Hardy Weinberg equilibrium are frequently used in population studies to test for random mating and for the presence of factors altering the allele frequencies. Based on the observed allele frequencies in an equilibrium population, the genotype frequencies will give an approximately fit to the Hardy-Weinberg formula. However, if the genotype frequencies deviate significantly from the expected Hardy-Weinberg values, it is necessary to investigate for evidence of:

1. Patterns of nonrandom mating, including inbreeding and crossbreeding or
2. A change in the allele frequencies due to selection, migration, mutation, or genetic drift

3.3 Inbreeding

Inbreeding is the production of offspring by related parents. The most extreme form of inbreeding is self-fertilization, which is possible in a few species used for aquaculture such as some species of scallop. Systematic patterns of mating, such as full-sib matings, half-sib matings, all result in inbreeding. In a population of finite size, inbreeding accumulates over time even with random mating as by chance some related individuals will be mated.

The inbreeding coefficient, F, is probability that two alleles at any locus in an individual are identical by descent (ie. both alleles can be traced back to a single common origin). It is possible to calculate the rate of accumulation of inbreeding, ΔF, both for systems of non-random mating and for a population of finite size with random mating. For example, the rate of accumulation of inbreeding per generation in a randomly mating population of finite size is:

$$\Delta F = 1/2N \tag{3.1}$$

where *N* is the effective number of breeding individuals, or more precisely the size of an idealized population. The idealized population is a population where (Falconer 1981):

- mating is restricted to members of the same line. In other words migration is excluded
- the generations are distinct and do not overlap
- the number of breeding individuals in each line is the same for all lines and in all generations. Breeding individuals are those that transmit genes to the next generation

- within each line, mating is random, including self-fertilisation in a random amount
- there is no selection at any stage
- mutation is disregarded

The effective population size of most aquaculture breeding populations tends to be small, as the high fecundity of aquaculture species means that relatively few broodstock are required to breed large populations. For example, with N_m males and N_f females used each generation, the effective population size is:

$$N_e = \frac{4N_m N_f}{N_m + N_f} \tag{3.2}$$

So if a fish breeding scheme uses a mating ration of one male to two females, and 25 males are used each generation (50 females), the effective population size is approximately 67. The rate of accumulation of inbreeding in this breeding program, using the formula above, would be 0.75% per generation.

The effect of inbreeding is to increase the frequency of the homozygote genotypes. One direct consequence of this is that more recessive alleles are expressed in inbred populations. Most deleterious or disease conditions are recessive, and are therefore more likely to be expressed in inbred populations. So inbred populations therefore suffer from a reduction in fitness. Inbreeding can also lead to inbreeding depression, which will be discussed in Chapter 6.

3.4 Crossbreeding

In contrast to inbreeding, which increases the frequency of homozygotes, crossbreeding or the mating of unrelated individuals increases heterozygosity. By increasing the frequency of heterozygous genotypes, recessive alleles are less likely to be expressed, and so the fitness of the population can be increased (for one generation following crossing at least). Increased heterozygosity of crossbred individuals is often observed as hybrid vigor or heterosis, when the average value of the offspring for a particular trait exceeds the mean of the average values of the two parental lines. Heterosis is exploited in some plant and animal breeding systems, but is not widely used in aquaculture species. An exception is Nile-tilapia in the Philippines. In a large scale trial, two different crosses, *O. aureus* x *O. spilurus* and *O. mossambicus* x *O. niloticu*s showed 22 and 25% percent heterosis respectively for body weight. Another example is that hybrids of *O. mossambicus* and *O. niloticus* have higher tolerance to salinity than pure *O. niloticus* (Tayamen et al., 2002).

In contrast to genotype frequencies, allele frequencies in the population are not affected by inbreeding or crossbreeding *per se*. Other forces, such as natural or artificial selection, mutations, genetic drift, and migration, which together are responsible for evolution, do alter frequencies of alleles in the population.

3.5 Selection

Selection occurs when particular genotypes have greater reproductive success than other genotypes. For example, if a particular allele at a gene confers a reproductive advantage over other animals in the population, then the frequency of the allele in the population could be expected to increase over time.

The effect of selection can be demonstrated considering a large random mating population with a single locus, with two alleles *A* and *a* in the population (from Gardner and Snustad, 1981). Let 1-q be the frequency of *A*, and q the frequency of *a*. *A* is completely dominant in this case. A constant selective advantage is assumed for the dominant phenotypes (*AA* and *Aa*) over the *aa* genotype. Reproductive rates for genotypes *AA*, *Aa* and *aa* are 1:1:1-s respectively. The selection coefficient s, measures the disadvantage of an organism, relative to a particular reference genotype under selection, and consequently represents the intensity of selection. If *aa* is lethal (for example, Albinism in the channel catfish) then s = 1. If q in the initial population and s are both known, the frequency of *a* in each successive generation can be calculated as

$$\Delta q_s = -\frac{sq^2(1-q)}{1-sq^2} \tag{3.3}$$

This formula can be used to show the many generations are required to reduce the frequency of lethal recessive alleles (in the case where heterozygotes *Aa* are not affected). In the example of the albinism allele above, if the initial allele frequency of *a* is 50%, it will take approximately 50 generations to reduce the frequency of the *a* allele

Table 3.1 Frequencies of genotypes and genes when selecting against recessive genotypes (*aa*). The homozygote dominant (*AA*) and heterozygote (*Aa*) Phenotypes can not be distinguished. q is frequency of the recessive gene *a*.

Generation	% genotype			Gene frequency
	AA	*Aa*	*aa*	of q
0	25.00	50.00	25.00	0.500
1	44.44	44.44	11.11	0.333
2	56.25	37.50	6.25	0.25
3	64.00	32.00	4.00	0.20
4	69.44	27.78	2.78	0.167
5	73.47	24.49	2.04	0.143
10	84.03	15.28	0.69	0.083
50	96.19	3.77	0.04	0.049
100	98.05	1.94	0.01	0.010

to 5 % in the population as shown in Table 3.1. Selection against the *a* allele is rapid at first, but decreases as the frequency of *a* declines. In fact, the *a* allele may never be

entirely eliminated from the population through selection, as when q is small enough to allow homozygotes to occur very infrequently, the recessive allele is still harboured in *Aa* heterozygotes. In the case where heterozygotes are more viable than both homozygotes, neither allele will eliminate the other since the superior heterozygotes contain both alleles. This phenomenon is referred to as balanced polymorphism.

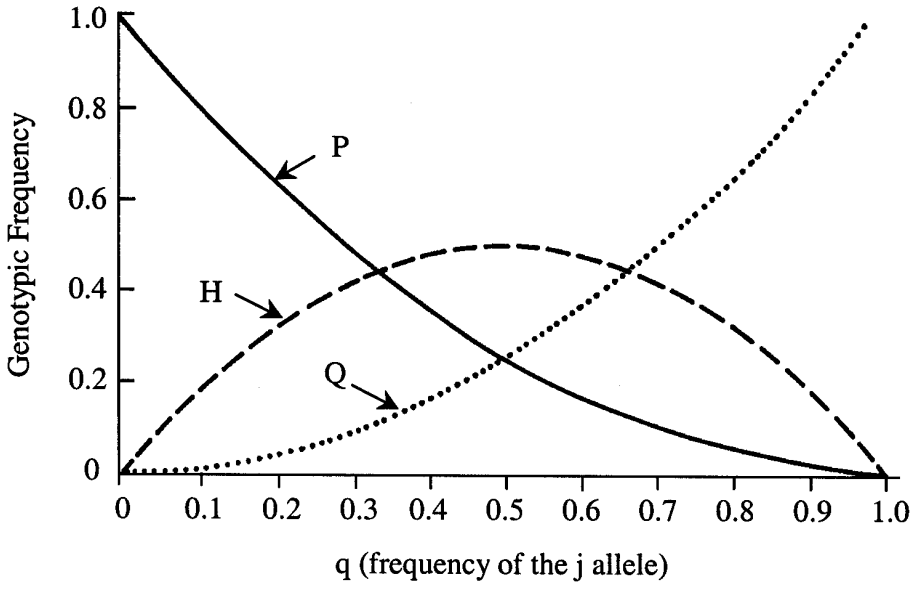


Figure 3.2. Relationships between allelic and genotypic frequencies shows the P frequency of homozygote dominant individuals, H frequency of heterozygotes, and Q frequency of homozygote recessives.

In nature, the selection coefficients at any one locus are generally very small, and so natural selection needs many generations to alter allele frequencies within a population. This allows genetic variation in the form of different alleles to remain within the population. As a consequence, the population may be able to respond to changes in the natural environment, if successive changes favor different alleles.

The consequence of artificially selectively breeding animals is also a change in allele frequencies of the genes which affect the selected traits are changed over time. If the trait is determined by one or a small number of genes, for example skin colour in tilapia, then the change in allele frequency due to intense selection can be rapid. This can result in rapid changes in the proportions of the genotypes, Figure 3.2.

3.6 Mutation

Mutations change one allele to another, for example *A* to *a*. Neglecting backward mutation (ie *a* to *A*), the change in frequency of *a* is proportional to the mutation rate (u)

and frequency (p) of *A*. The number of generations (t) necessary for a given change in allele frequency can be computed from the equation (Falconer, 1981):

$$p_t/p_o = (1-u)^t \tag{3.4}$$

where p_t and p_o is the allele frequency in generation t and the base generation, respectively. With small per locus mutation rates, a very large number of generations are required for mutation to have an appreciable effect on allele frequencies. For example, 9050 generations would be required to change the frequency of *a* from 0.1% to 1 % by accumulation of mutations at a mutation rate of 10^{-6}.

A more important consequence of mutation is the creation of new allelic variants, which can be acted upon by selection. The accumulation of changes in DNA over millions of years has provided the source of variation that makes one species distinct from another, and also underlies the differences between individuals of the same species. Whereas some mutations occasionally can benefit the organism, the majority of mutations is detrimental, and are generally kept at low frequency in the population through selection. It is possible to predict the trend in allele frequencies in the population considering both mutation and selection (s = selection intensity) simultaneously. If the frequency of the rare (and deleterious) allele is small, the change in allele frequency each generation is approximately $-sq^2$, while mutation increases the frequency of the rare allele (*a*) by u each generation. At equilibrium, both forces will be equal, and the frequency, which the rare allele will attain, $\hat{q}$ can be estimated as (Falconer, 1981):

$$\begin{aligned} -sq^2 &= u \\ \hat{q} &= \sqrt{u/s} \end{aligned} \tag{3.5}$$

So consider again our example of Albinism in catfish, where the *aa* genotype is lethal (s=1). If the mutation rate from the wild type to the albino allele is 10^{-6} each generation, and s=1, the estimated frequency which the allele will obtain in the population is 10^{-3}.

3.7 Genetic drift

Genetic drift is the random fluctuation in allele frequencies, which results from the sampling of alleles during breeding in small populations. The smaller the population, the greater this sampling effect will be. An extreme example is two parents both with heterozygous genotypes. If they only produce two offspring, there is a 25% chance both will be the *AA* genotype, and the *a* allele will be lost from the population. The process of genetic drift is important where the number of breeding individuals is in the tens or hundreds. Falconer (1981) noted that the direction of genetic drift could not be predicted, but it's magnitude, expressed as the variance in change of allele frequency as a result of sampling in one generation, can be stated as:

$$\sigma_{\Delta q} = q_0(1-q_0)/2N \tag{3.6}$$

where q_0 is the frequency of an allele in the base population, and N is the number of breeding individuals sampled each generation. With small numbers of breeding individuals sampled each generation, the variance in the change of allele frequency will be large, Figure 3.3.

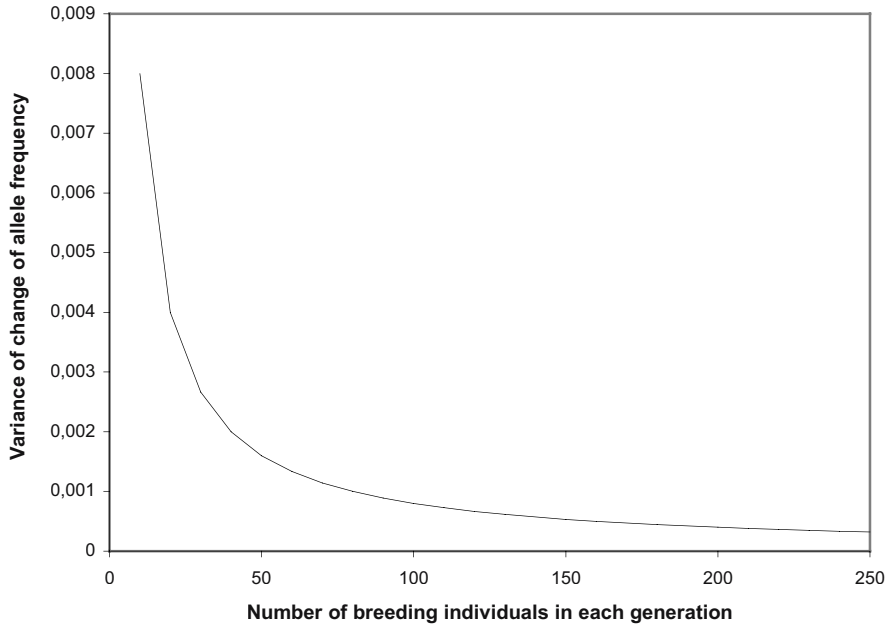


Figure 3.3. Variance of the change in gene frequency due to random drift with different numbers of breeding individuals in each generation.

The effect of genetic drift is to reduce genetic variability. This reduced variability may lower the capacity of a population to adapt to changes in the environment. Pronounced genetic drift is especially likely when a small group of individuals separate from a larger population. By chance these pioneers may carry only a fraction of the genetic variability of the parental population. Thus, the gene pools of the two populations are different at the very beginning. This phenomenon is called the founder effect. Genetic drift may also occur when many members of a population die, and the few remaining individuals mate to restore the numbers of the population. This process is described as a population bottleneck. Following a bottleneck, the gene pool of the new population will be restricted compared to the original pool.

3.8 Migration

Populations may become isolated from other populations of the same species by a number of mechanisms, including physical barriers such as rivers and mountain ranges.

The processes of genetic drift and selection in different environments can cause differences in allele frequencies between the populations. If there is then some movement of individuals between these otherwise isolated populations, and these immigrants interbreed with the native population, the allele frequencies in the receiving population may be altered. The frequency changes are proportional to the difference in allele frequencies between the populations and the rate of migration. Let us consider that a population consists of a portion of n new immigrants in each generation, and 1 – n individuals being natives. Let the frequency of a certain allele at a particular gene be q_n among the immigrants and q_o among the natives. Then the frequency of the allele in the mixed population, q_m will be:

$$q_m = nq_n + (1 - n)q_o = n(q_n - q_o) + q_o \tag{3.7}$$

Often, immigrants will come from neighboring populations, which may not have had time to develop markedly different allele frequencies from the population being considered. However, the import of animals from a remote population, or the introduction of sperm and eggs in domesticated population may dramatically alter allele frequencies, or indeed contribute entirely new alleles. This could be the case if some of the introduced alleles are adaptively superior in the new environment. The extreme case of such exchange occurs during interspecies hybridization (interbreeding between populations that are different enough to be ranked as separate species). Although much less common than gene flow within a species, the effects are more pronounced, as the allele frequencies may have diverged for a much longer time period. Indeed, many alleles in one species may not be present in the other (Gardner and Snustad, 1981).

3.9 Genetic distance and population divergence

As a result of the forces of genetic drift, selection and mutation, populations which are isolated from each other, can experience divergence in allele frequencies, particularly if population sizes are small or if selection is for different characteristics in the different populations (for example, different diseases). In fact the extent of the divergence in allele frequencies between populations can be used to infer how long the populations have been isolated and the extent of gene flow between them.

A simple statistical test for measuring the accumulated allele differences between multiple populations was developed by Nei (1975). Nei's genetic distance (D) is estimated as follows: Let x_i and y_i be the frequencies of allele A_i in two randomly mating populations X and Y. Drawing one allele at random from each population the probability for choosing identical alleles $j_{xy} = \Sigma x_i y_i$. The probability of drawing two identical alleles in population X is $j_x = \Sigma x_i^2$, and $j_y = \Sigma y_i^2$ in population Y. If J_{xy}, J_x and J_y are the arithmetic means of j_{xy}, j_x and j_y, respectively, over all loci in the genome, then the genetic distance D is given by the formula:

$$D = - \log_e I \tag{3.8}$$

where $I = J_{xy}/(J_xJ_y)1/2$.

Nei's genetic distance has been widely used to quantify genetic differences between populations. Recent advances in molecular techniques and the identification of highly polymorphic repetitive DNA elements have made it relatively easy and cheap to genotype a large number of loci, improving the estimates of genetic distance. (see Chapter 20).

This method was applied to study the genetic distance between eight tilapia strains (Macaranas et al., 1995). Farmed tilapia strains in the Philippines derived from introductions via Israel, Thailand, Singapore and Taiwan from 1972 to 1988 together with four wild-caught African strains from Egypt, Ghana, Kenya and Senegal were included. Genotypes at 30 loci were compiled and analyzed to study genetic distance.

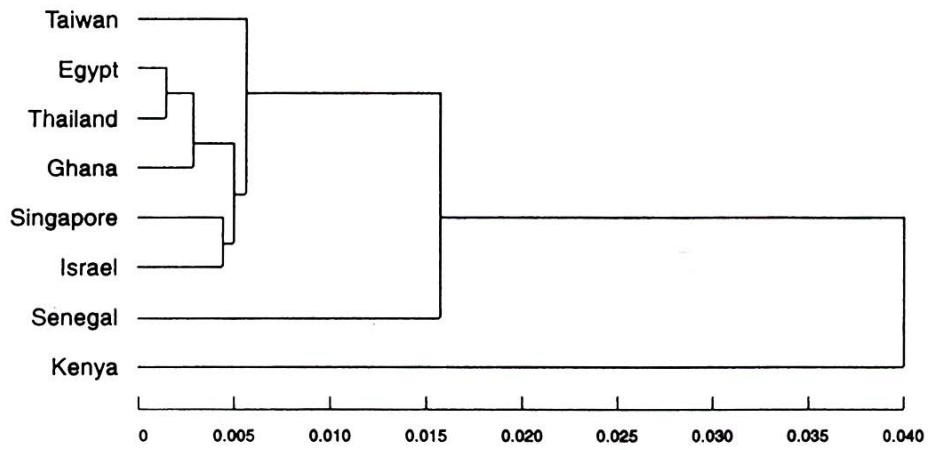


Figure 3.4 Dendrogram showing genetic relationships among the eight *Oreochromis niloticus* strains constructed from Nei's genetic distance values. Reproduced from Macaranas et al. (1995) by permission of Springer Science and Business Media.

The result is shown in Figure 3.4. The genetic distance among the four farmed Philippine strains were low and close to the strains from Egypt and Ghana while the Senegal strain and in particular the Kenya strain were widely distant apart.

3.10 Summary

By extending Mendelian laws from the level of the individual to the level of the population, it is possible to derive expressions for the changes in allele frequencies, and genotypes, at a particular locus in the population. The simplest scenario is random mating in the absence of selection, mutation, migration and genetic drift. The Hardy-Weinberg theorem predicts, that in this scenario, the genotype frequencies will be constant over generations. Non-random patterns of mating, such as inbreeding or

crossbreeding, alter the genotype frequencies, but not the frequencies of the alleles. The forces of selection, mutation, migration and genetic drift act to alter allele frequencies in the population. Such changes in allele frequency can cause divergence between two populations, which have the same ancestral gene pool.

4. POLYGENIC INHERITANCE

HANS BERNHARD BENTSEN

4.1 Introduction

The association between the Mendelian inheritance of alleles and changes in allele frequencies in populations (population genetics) is obvious, as described in Chapter 3. The relationship between allelic inheritance and the inheritance of complex traits like growth, maturity, behavior etc. (quantitative genetics) is often less apparent. Population genetics deals with allelic variation, which may often be detected by chemical analysis at the level of protein or DNA polymorphism. The effects of the alleles on the phenotype is usually unknown, or mainly of interest as markers of allelic variation. The analysis of the inheritance of complex traits, usually termed quantitative genetics, deals with phenotypic measurements and observations, and the underlying allelic constitution is normally unknown. Nevertheless, quantitative genetics is also based on the Mendelian laws of inheritance.

A simple example of the relationship between Mendelian genetics, population genetics and quantitative genetics is shown in Table 4.1. A gene occurs in a population as two different alleles, *a* and *A*, both with an allele frequency of 0.5. At Hardy-Weinberg equilibrium (Chapter 3.2), this will result in a distribution of allelic genotypes as shown in the table. If the alleles have an effect on a phenotypic trait, e.g. that *A* has an additive effect of +1 unit of measurement compared to *a*, the genetic effects of the different allelic genotypes on the phenotypes will be as shown in the table.

Table 4.1. The frequency of different allelic genotypes and genetic effects on the phenotypes in a population with alleles *a* and *A* at frequencies of 0.5 and an additive genetic effect of *A* relative to *a* of +1.

Frequency	Allelic genotypes	Genetic effect
25%	*aa*	0
50%	*aA* or *Aa*	+1
25%	*AA*	+2

Assuming that other genetic or non-genetic effects do not affect the trait, the population may now be characterized by numerical statistics, e.g. by the mean genetic effect of the gene on the phenotypic trait (=1.0) and by the genetic variance (=0.5) or the genetic standard deviation (=0.71) (may easily be confirmed by applying the standard formulas). These parameters may be obtained from phenotypic measurements without any knowledge about the underlying allelic variation.

4.2 Limitations of the single gene model

A single gene may only generate a limited number of genetic effects on a phenotypic trait, depending on the number of different allelic forms that the gene may occur in.

T. Gjedrem (ed.), Selection and Breeding Programs in Aquaculture, 35–43.
© 2005 *Springer. Printed in the Netherlands*

Polymorphism studies usually fail to detect more than two to five alleles of a given gene, at least at the protein level of expression. As shown in the example above, two alleles may only generate three different genetic effects (or only two in the case of complete dominance of one of the alleles). Unless one of the alleles is very rare, all possible genetic effects will be expressed among the individuals in the population. If the effects of the alleles are additive, selection may then only increase the occurrence of the most desired genetic effect in the population until the desired allele is fixed. This will usually not generate individuals with a genetic effect on the trait that is better than the best performing individuals in the base population before selection was started. The fixation of the genetic variation for the trait will also prevent further response to selection. In other words, it is then expected that the best performing individuals in the base population represent the selection limit for the trait.

This is not what is observed in most selection experiments or breeding programs. A selection experiment for increased pupa weight in *Tribolium* showed no evidence of reduced genetic variation after 119 generations of selection (Enfield, 1979). A difference of 28 genetic standard deviations had then been established between the control (representing the base population) and the selected populations. Bearing in mind that the pupa weight of practically all individuals in the base population is expected to fall within the range of three genetic standard deviations below or above the mean value (according to the normal distribution), all individuals in the selected population were considerably heavier than the heaviest individuals in the base population. Modern Scandinavian broiler chicken breeding programmes have reduced the growth period to market size (1.4 kg) from 10-12 weeks to 5 weeks during the last 25 years (Sørensen, 1986), and response to selection is still maintained. The wild chicken progenitors normally do not reach broiler market size even as adults. How can such results be explained?

One possible explanation may be new mutations. Falconer and Mackay (1996) stated that new variation from mutated genes may begin to contribute to the response after perhaps 20 generations, and that all genes segregating in the base population should have been brought to fixation after some further time. Continued response will then depend entirely on the mutations that have accumulated during the selection. On the other hand, Enfield (1988) summarized the common assumption that the effective population size has been too small and the number of generations too few in most laboratory experiments and applied animal breeding programmes for new mutations to play a significant role in selection response. Mutations have generally been assumed to occur at rates of 10^{-5} to 10^{-6} per generation, and in most cases with adverse effects. Certainly, the long domestication history of most farm animals allows for mutations to play a role in the genetic differentiation between the farmed populations and their wild progenitors. However, a large proportion of the current breeds and varieties have been formed during the last 150 years, and a major proportion of the additive genetic improvement has happened during the last 50 years.

Another explanation to the observed sustainability of response and genetic variability during long-term selection is that quantitative traits are probably influenced by a large

number of segregating genes, each with a relatively minor impact on the trait. The implications of this polygenic inheritance will be addressed in the present chapter.

4.3 A simplified polygenic model

A simplified example may illustrate the polygenic relationship between alleles, allelic genotypes and genetic effects. Figure 4.1 shows the allelic genotypes and the distribution of genetic effects for a trait coded by two independently segregating genes when each gene has two allelic forms with allele frequencies of 0.5, one zero (i.e. background) allele and one allele adding one unit of measurement to the trait (a plus allele).

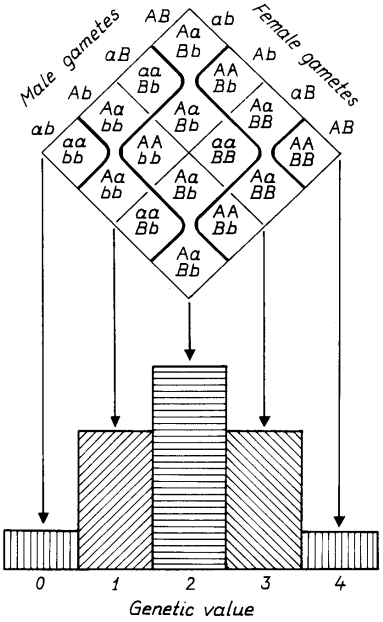


Figure 4.1. The allelic genotypes and the distribution of genetic effects (genetic values) of a trait equally affected by two loci, each with a zero allele (*a* and *b*) and an allele adding one unit of measurement to the trait (*A* and *B*) at frequencies of 0.5 in the parent generation.

It should be noted that any genetic effect except the effects of the extreme, fixed genotypes in Figure 4.1 might now be obtained by several allelic combinations. Consequently, a group of individuals sharing a given genetic effect other than the extremes may carry all the alleles of the population, only at different frequencies.

In principle, the approach used to construct Figure 4.1 may be applied to any number of genes and alleles and any relative effect of the plus alleles. As long as it is assumed that the genes are segregating independently and that each gene has two allelic forms with

the same effect of the plus allele on the trait and a common frequency of all plus alleles, the distribution of genetic effects will follow a binomial distribution with the parameters *n* (number of repetitions) equal to the number of alleles or twice the number of genes, and *p* and *q* (the probabilities of the two outcomes) equal to the frequency of the plus and the zero alleles. Figure 4.2 shows the distribution of genetic effects for a trait coded by 5 genes (10 alleles) and 10 genes (20 alleles) under these simplified assumptions and an allele frequency of the plus alleles of 0.5. The standard normal distribution curve is also shown.

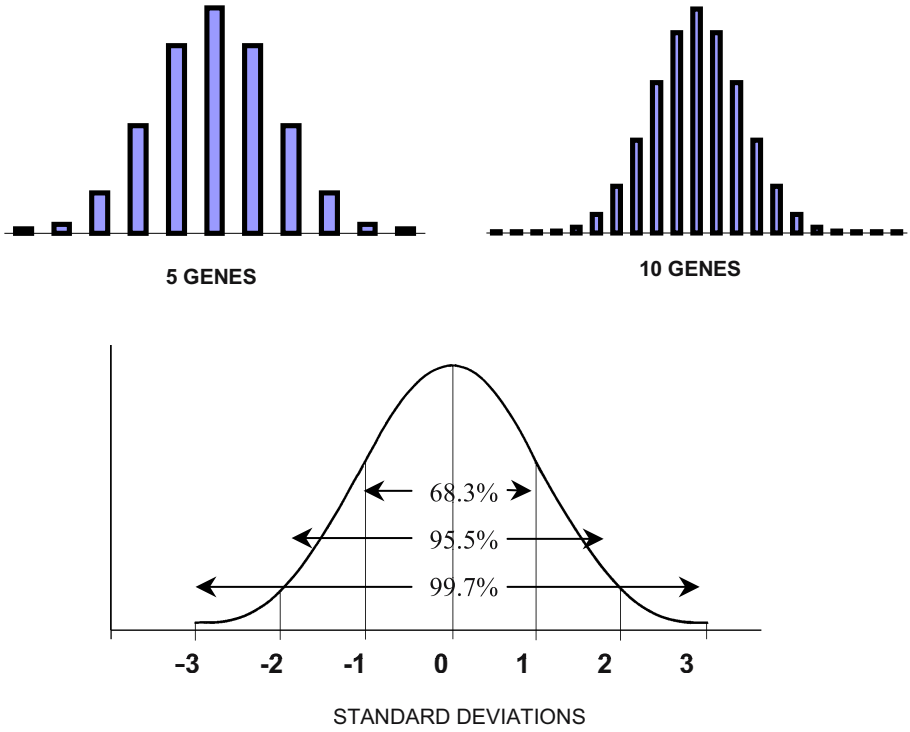


Figure 4.2. The distribution of genetic effects in a population for a trait coded by 5 or by 10 genes under the same simplified assumptions as in Figure 4.1, compared to the distribution according to the standard normal distribution curve.

From Figure 4.2 it can be seen how the number of classes of the genetic effect increases as the number of genes affecting the trait gets larger. The figure also illustrates that the binomial distribution of the genetic effects in the population approaches a standard normal distribution as the number of genes increases, at least when *p* and *q* are not very close to zero or one (i.e. fixation or loss of the alleles). The mean value of the binominal distribution will be *np* and the variance *npq*. As in Figure 4.1, all classes except the

extreme fixed effects may carry all the alleles of the population. Figure 4.2 also shows that the frequency of the extreme fixed allelic genotypes is reduced as the number of genes increases. The frequency of the extreme genetic effects will be 9.8×10^{-4} for 5 genes and 9.5×10^{-7} for 10 genes.

As shown in the figure, the common presentation of the standard normal distribution curve is open in the ends and does not include the extreme values. In fact, the normal distribution curve continues indefinitely in both ends, but the frequencies of values outside three standard deviations from the mean are rare (all together 0.3%), and increasingly rare for the more extreme values. The span of the binomial distribution is, however, limited between the outcomes of zero plus alleles (resulting in the combined background effect of all the zero alleles) and *n* plus alleles (adding *n* plus allele effects to the background effect). Figure 4.3 shows the distribution and span of genetic effects measured in genetic standard deviations according to the simplified polygenic model when the number of genes (loci) affecting the trait is 10, 100 or 300.

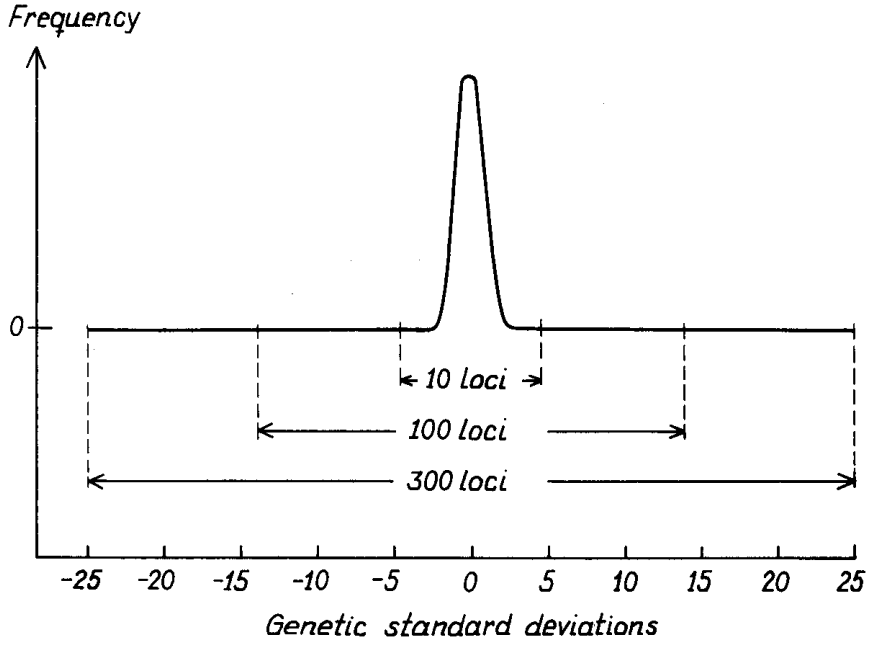


Figure 4.3. The distribution of genetic effects and the theoretical genetic span measured in genetic standard deviations for different numbers of genes (loci) affecting the trait.

According to the normal distribution, 99.7% of the population is expected to deviate less than three standard deviations from the population mean, and the probability of deviations larger than four standard deviations is practically zero. This means that only a limited interval of the possible range of genetic values will be expressed in a real population when the number of loci affecting the trait is above a certain minimum. The frequency of the fixed extreme genotypes in Fig. 4.3 will be 2.4×10^{-191} for 300 genes.

4.4. How many genes?

A crucial point for the conclusions until now has been the number of loci affecting the trait. Traits coded by a single or a few gene(s) are well known because they are easy to detect (e.g. colour traits). In the early period of quantitative genetics, it was frequently assumed that a limited number of loci could be identified to explain completely the genetic variability of more complex traits, and that selection limits and fixation would be reached quite rapidly (see e.g. Lerner, 1950). Later experience has not confirmed these early expectations, and in several more recent selection experiments, the minimum estimates of the number of loci involved in some complex traits have been close to or above 100 or even above 300 (Enfield, 1974; Falconer and Mackay, 1996). Other observations have given lower minimum estimates, but it is uncertain if this is due to the nature of the trait selected for or due to the populations studied. In populations with a long-lasting, closed history, a large proportion of the genes normally causing variation in the trait may have been fixed because of genetic drift or inbreeding.

Such minimum estimates are based on several assumptions and simplifications (see e.g. Lande, 1981; Falconer and Mackay, 1996), and may not be regarded as final evidence about the possible number of loci affecting a quantitative trait. However, it may easily be argued that a large number of metabolic proteins are likely to be involved in complex processes like, for example, growth or reproduction, and that many of these proteins are direct or indirect products of variable genes. Recent research on the human genome has suggested that the total number of protein-coding genes is about 30 000 to 40 000. Some fish species may have even higher numbers.

Furthermore, studies of the molecular genetics of DNA (see e.g. Singer & Berg, 1991) have identified genetic elements that may act as additive genes in a quantitative context without coding for metabolic proteins. Regulatory DNA sequences and sequences coding for DNA regulatory proteins have been shown to determine the amount, timing and location of gene expression. Genetic differences involving regulatory mechanisms have been shown to affect gene expression and would consequently be expected to cause variation in additive genetic values as if they were regular alleles. Each protein-coding sequence may be influenced by several such regulatory mechanisms, and regulatory interactions occur across the genome. The number of loci that determines the quantitative genetic variation of a trait may consequently be much larger than the number of genetically coded, variable proteins directly involved in the expression of the trait. In this context, even a number of loci as high as 300 may not be unreasonable for a complex, polygenic trait.

4.5. Implications for the response to selection

The effect of different frequencies of the plus alleles on the genetic difference between populations according to the simplified polygenic model is shown in Figure 4.4 for a trait affected by 300 genes.

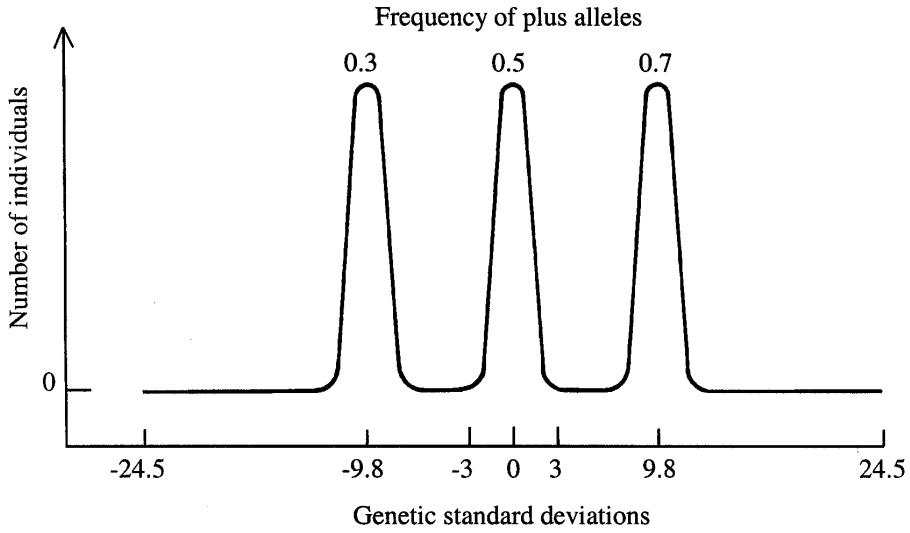


Figure 4.4. The distribution of genetic effects, measured in genetic standard deviations according to the simplified polygenic model, for a trait affected by 300 genes when the frequency of all plus alleles is 0.3, 0.5 or 0.7

It is evident from the figure that even moderate differences in allele frequencies may generate genetically non-overlapping populations for a quantitative trait when the number of loci affecting the trait is large. Even intensive selection experiments will normally not produce a selection response (a change of the population mean) that is much larger than 0.5 genetic standard deviations per generation. Applied selection programs to improve genetically complex traits may consequently generate continuous response to selection over a large number of generations without other genetic consequences than moderate changes in allele frequencies, given that the trait(s) selected for are coded by a large number of genes, each with a minor effect.

In fact, the possible differences created by simultaneous, moderate changes in a large number of alleles may easily become larger than what may seem biologically feasible. For example, the genetic standard deviation for adult body weight in farmed Atlantic salmon (mean value about 4 kg) has often been estimated to be about 1 kg. The distance between the extreme fixed values in Figure 4.4 will then be 50 kg for salmon of the same age and under the same feeding and management regime. Bentsen (1994) has given other examples. It is likely that biological restrictions will limit the selection response long before such differences may be realized, i.e. before the extreme, fixed genetic values are reached. This agrees well with experimental findings (e.g. the selection experiment for increased pupa weight in *Tribolium* reported by Enfield, 1979) here response to selection seems to slow down after an extensive number of generations with steady genetic progress, but without any signs of significant loss in genetic variation for the trait under selection.

From the formula for the variance of the binominal distribution (*npq*), it may easily be confirmed that the maximum variance occurs at allele frequencies of 0.5, but that a change in the frequencies of the plus alleles to 0.2 or 0.8 will only reduce the standard deviation of the distribution curve with 20%, and a change to 0.1 or 0.9 will cause a reduction of 40%. Consequently, only extreme allele frequencies will cause severe reductions of the magnitude of the standard deviation of the distribution of genetic effects. From Figure 4.4 it may also be seen that considerable response to selection may then occur without large effects on the standard deviation when the number of genes affecting the trait is large.

4.6. Limitations of the simplified polygenic model

The assumptions of the simplified polygenic model are: (1) that all variable genes have two possible alleles, (2) that the effect of the plus allele of all genes is the same, (3) that all plus alleles occur at the same frequency, (4) that the genes are segregating independently (no linkage) and (5) that all effects of the alleles are additive (no dominance or epistatic effects). This will of course rarely be the case in a real situation. What are the consequences when these assumptions are violated?

(1) Variable genes may often have more than two possible alleles. Some gene regions like the Major Histocompatibility Complex are characterized by an extensive allelic variation that seems to be crucial for it's biological function. However, most genes seem to occur in a rather limited number of allelic forms. The simplified model will still be quite valid if the variation caused by the gene is mainly due to one allele having a different effect on the trait compared to several other alleles with a more similar effect. If a gene has more than two distinctly different allelic effects on the trait, this will increase greatly the number of classes of possible genetic effects in the population, but the main conclusions from the simplified model is not likely to by severely affected.

(2) All variable genes will probably not have an equal effect on the trait. Some of the genes may be genes (or complexes of closely linked genes) with a major effect on the trait. A meta-analysis across livestock data suggest that most of the genes affecting quantitative traits are expected to have relatively small effects, and that genes with large effects occur in much lower numbers (Hayes and Goddard, 2001). Furthermore, major genes with large additive effects are expected to reach fixation quite fast during selection (Sehested & Mao, 1992). The remaining genetic variation for traits that have been subject to directional selection would then tend to be caused mainly by a large number of minor genes. However, many aquaculture populations are still genetically wild and have not been subject to significant directional selection for production traits. Major genes that are still not fixed may then magnify the additive genetic variation and increase the response to selection during some initial generations.

(3) All plus alleles are not expected to occur at the same frequency. A dispersion of the frequencies of the plus alleles around the assumed average frequency may often reduce the variation (and standard deviation) of genetic effects in the population, especially if the dispersion is large. A reduced genetic standard deviation will imply that the distance

between the theoretical, extreme fixed genetic effects will increase when compared to the variation within the population. Selection will bring plus alleles that occur at high frequencies more rapidly towards fixation, and this will tend to reduce the variation of genetic effects in the population. On the other hand, selection will also increase the frequency of rare plus alleles, and this will tend to increase the variation. All together, the variation of genetic effects will then often be quite unchanged by selection, as is the case for the simplified model.

(4) When the number of genes affecting a trait is large, it is evident that all genes cannot segregate independently, since the genes will be grouped on a limited number of chromosomes with a certain degree of linkage between them. Genes that are closely linked will tend to be inherited as a single gene that affects the trait with their combined effect. More distantly linked genes will recombine over time and the effects will be as described for the simplified model, but with some delay.

(5) Large non-additive genetic effects may disturb several of the conclusions drawn from the simplified polygenic model. Dominance may reduce the number of classes of genetic effects in the population, while epistatic effects may increase the number of classes. Both dominance and epistatic effects may reduce the response to selection because of recombination of interacting alleles during the inheritance from one generation to the next.

4.7. Summary

The inheritance of allelic forms of a single gene cannot explain the large and rapid changes in complex genetic traits that can be induced by selection in a population. Selection on a single gene will normally not be able to generate genetic effects outside the effects observed in the base population before selection, and fixation of the desired allele will soon lead to extinction of the genetic variation. This does not agree with observations from selection on complex traits. However, by extending the Mendelian laws of inheritance to a simplified model of polygenic inheritance, it is possible to explain how a continuous response to selection may be observed far beyond the genetic values observed in the base population, and without loss of variation in the genetic effects on the trait. The model considers the simultaneous additive genetic effect of many segregating genes, each with a minor effect on the trait. If the number of genes affecting the trait is sufficiently large, selection limits may be caused by biological constraints rather than fixation of the desired alleles.

5. BASIC STATISTICAL PARAMETERS

TRYGVE GJEDREM AND INGRID OLESEN

This is a brief introduction to the basic parameters in common use in animal breeding. We will not show the biometrical relationships and deduce the formula for the quantitative parameters, because such information can be found in many other textbooks and papers. For interested readers we recommend: Introduction to quantitative genetics by Falconer and Mackay (1996), The genetics of populations by Lush (1994), and Statistical methods by Snedecor and Cochran (1980).

5.1 Variation in traits

5.1.1 MEAN, STANDARD DEVIATION AND VARIANCE

To characterise a trait/character for breeding purposes, at least the mean or average value and standard deviation of the trait must be known. In Table 5.1 the mean, standard deviation, and coefficient of variation are given for some production traits in aquaculture species. In addition, it is necessary to know the distribution of the trait when deciding which statistical method shall be used for further study of the variation in traits.

The average value of a sample, $\overline{x}$, of such a trait is:

$$\overline{x} = \sum_{i=1}^{N} x_i / N \tag{5.1}$$

where $\sum_{i=1}^{N} x_i$ is the sum of the value of the trait X for all N individuals in the sample.

The variance provides information on the variation among animals in a particlar trait. The variance of N observations in a sample, $\hat{\sigma}^2$, can be estimated as:

$$\hat{\sigma}^2 = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \overline{x})^2 = \frac{1}{N-1}\left[\sum_{i=1}^{N} x_i^2 - \frac{1}{N}\left(\sum_{i=1}^{N} x_i\right)^2\right] \tag{5.2}$$

The variance is a measure of the average squared difference between each observation and the mean. The variance is measured in squared units, such as g^2. The standard deviation is the square root of the variance, and is often used as the measure of the variation between measurements. It can be more useful than the variance, because the dimensions of the mean and the standard deviation are the same. The standard deviation of N observations in a sample is estimated as:

T. Gjedrem (ed.), Selection and Breeding Programs in Aquaculture, 45–72.
© 2005 *Springer. Printed in the Netherlands.*

$$\hat{\sigma} = \sqrt{[\frac{1}{N-1}(\sum_{i=1}^{N} x_i - \overline{x})^2]} \quad (5.3)$$

Table 5.1 Average, standard deviation and coefficient of variation for different traits in different species.

	Average	Standard deviation	Coefficient of variation
Atlantic salmon:			
Body weight of fingerlings, g	9,8	7,6	80
Length of fingerlings, cm	8,9	2,3	26
Body weight, kg	6,6	1,89	29
Gutted weight, kg	6,2	1,79	29
Length, cm	81,2	6,4	8
Condition factor	1,2	0,14	12
Dressing percentage	93,9	0,8	-
Fat in filet	16,0	2,5	16
Rainbow trout:			
Body weight of fingerlings, g	13,3	4,4	33
Length of fingerlings, cm	3,4	0,73	21
Body weight, kg	3.0	0,74	25
Gutted weight, kg	2.3	0.59	26
Length, cm	59,0	4,4	8
Condition factor	1,6	0,15	10
Dressing percentage	87,3	1,9	-
Fat in filet	14,8	2,6	18
Rohu carp:			
Body weight of fingerlings, g	31.7	16.7	53
Body weight, g	401.1	109.4	27
Tilapia:			
Body weight of fingerlings, g	3.5	1.2	35
Body weight, g	72.6	25.2	35
Shrimp, P. vannamei:			
Stocking weight, g	1.5	0.7	44
Harvest weight, g	19.3	3.4	18

The variance has some important properties that are often used in animal breeding. The variance of trait X will be denoted by var(x) and α is a constant.

1 $var(X + \alpha) = var(X)$
Adding a constant to each observation does not change the variance, because the variation between observations remains the same, although the mean changes by the constant α.

2 $var(\alpha X) = \alpha^2 var(X)$
because $(\alpha X - \alpha \overline{X})^2 = \alpha^2 (xX - \overline{X})^2$, such that the variance increases by α^2

3 If X and Y are two independent traits, then
a) var(X + Y) = var(X) + var(Y)
b) var(X - Y) = var(X) + var(-Y) = var(X) +$(-1)^2$var(Y) = var(X) + var(Y)
c) $\text{var}(\sum_{i=1}^{N} x_i) = N\,\text{var}(X)$ from 3a) assuming equal variance for each X_i and all the X_i are independent

4 $\text{var}(\frac{1}{N}\sum_{i=1}^{N} x_i) = \frac{1}{N^2}\text{var}(\sum_{i=1}^{N} x_i) = \frac{1}{N}\text{var}(x)$
by combining properties 2 and 3 where 1/N is the constant. This means that the variance of a mean is $1/N^{th}$ of the variance of the observations. The variance of the mean of a sample of N observations is a measure of the precision of the mean. A mean with a smaller variance indicates a more precise estimate than a mean with a larger variance.

The square root of the variance of a sample mean is the standard error of the mean:

$$\sqrt{\text{var}(\bar{x})} = \sqrt{\frac{\text{var}(x)}{N}} = \text{standard error} \tag{5.4}$$

The standard error decreases as the number of observations increases.

5.1.2 COEFFICIENT OF VARIATION

A very useful statistical parameter for comparing variation across traits and species is the coefficient of variation (CV). While the average and standard deviation tend to change together, the coefficient of variation is independent from the mean. The CV is calculated as the standard deviation as percent of the mean:

$$CV = (\sigma / \bar{x})100$$

Some examples of coefficient of variation in body weight are given in Tables 5.1 and 5.3. Body weight has a very high coefficient of variation in fish and shellfish (CV = 17 - 29) compared to the coefficient of variation for body weight in farm animals like cattle, pigs and poultry, CV = 7-10 (Gjedrem, 1998). Body length, condition factor and fat percentage in fish have rather low values of the coefficient of variation however.

5.2 Distributions of observations

5.2.1 NORMAL DISTRIBUTIONS

In Figure 5.1, the distribution of a continuous trait is shown. This distribution is called a normal distribution, characteristic of many biological traits. Body weight of Atlantic salmon is used as an example in Figure 5.1. To get a smooth curve as in Figure 5.1 a large number of observations, $x_1, x_2, \ldots, x_N$, are required, where N is the number of observations or individuals recorded.

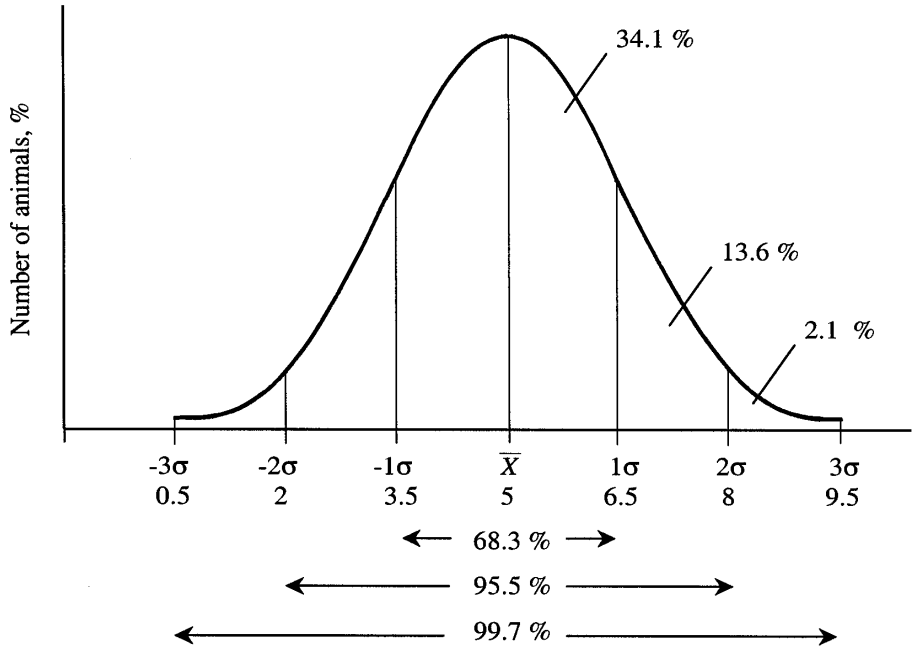


Figure 5.1 Distribution of body weight of 3 years old farmed Atlantic salmon with mean, X = 5.0 kg, and standard deviation, σ = 1.5 kg. Area under the curve and between ± 1σ represent 68.3%; ± 2 σ represent 95.5 % and ± 3 σ represent 99.7 % of the total number of observations.

The normal distribution is a bell-shaped curve and determined by two constants or parameters, the mean which locates the centre of the distribution, and the standard deviation which is a measure of the «spread» or the variation of the measurements, as shown in Figure 5.1. The curve is symmetric around the average, $\overline{X}$. The sum of all frequencies is equal to 1, and the area under the curve will therefore also equal 1. The frequency of observations between $\overline{X}$ and some point ($\overline{X}$ + f) is the same as the frequency of observations between $\overline{X}$ and ($\overline{X}$ - f). As shown in Figure 5.1 the normally distributed curve covers approximately 6σ, 3σ above and 3σ below the average, $\overline{X}$. In a normal distribution, 68.3 % of all observations are within one standard deviation above and below the mean, 95.5 % of the observations are within ±2σ, and 99.7 % of all observations are within ±3σ. Hence, a normally distributed trait is very well described when both the standard deviation and the average is known. Records of a normally distributed trait may deviate from normality if the sample of data is not random or when the data has been selected.

Many of the statistical techniques are based on the assumption that the traits are normally distributed. Therefore tests for normality may be useful. The test for skewness measures if the distribution is symmetrical around the average. A second test is the test for kurtosis, which measure if there is excess or a deficit of observations around the average. Sokal and Rolph (1981) give methods to test for both skewness and kurtosis.

Measurements of a trait which are not normally distributed, may become normal after a simple transformation. The square root and the logarithm functions are often used for such transformations.

A variable, X, which is normally distributed with a given mean μ and variance σ^2 is often given as $X \sim N(\mu,\sigma^2)$.

5.2.2 NON-NORMAL DISTRIBUTIONS

Sometimes observations of biological data are extremely skewed. In Figure 5.3 distribution of body weight of fingerlings of three salmonid species are shown. This skewed distribution can partly be explained by relatively small size differences during the early stages of the fish's life were magnified by the strong competion within population. Such hierarchical structure with a few big fish was also observed in a population of common carp by Moav and Wholfarth (1973).

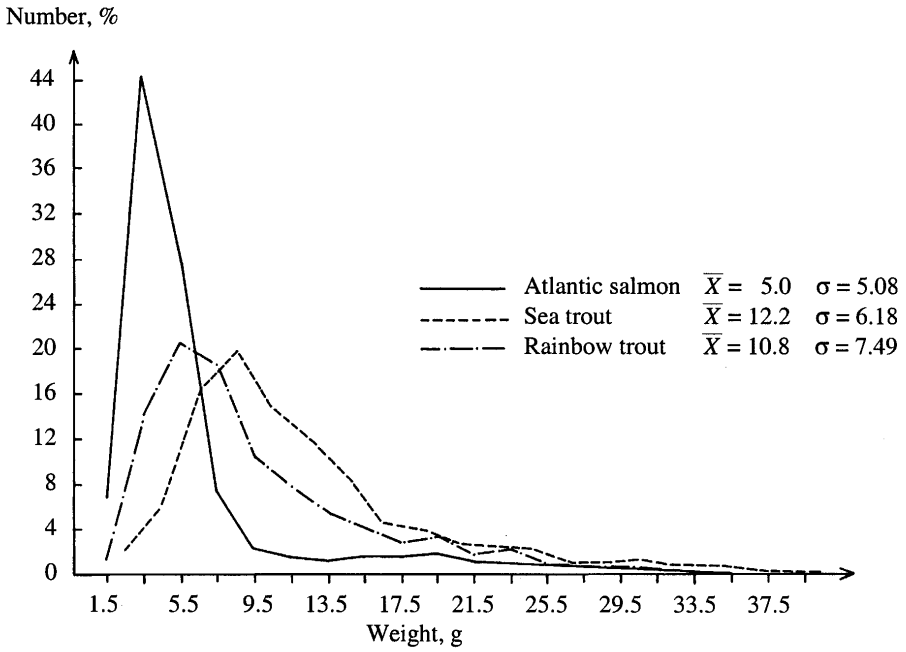


Figure 5.3 Distribution of body weight for fingerlings of different species of salmonids.

The distribution of a trait in the original population may be far from normal, but the distribution of sample means tends to become normal as the size of the samples increases.

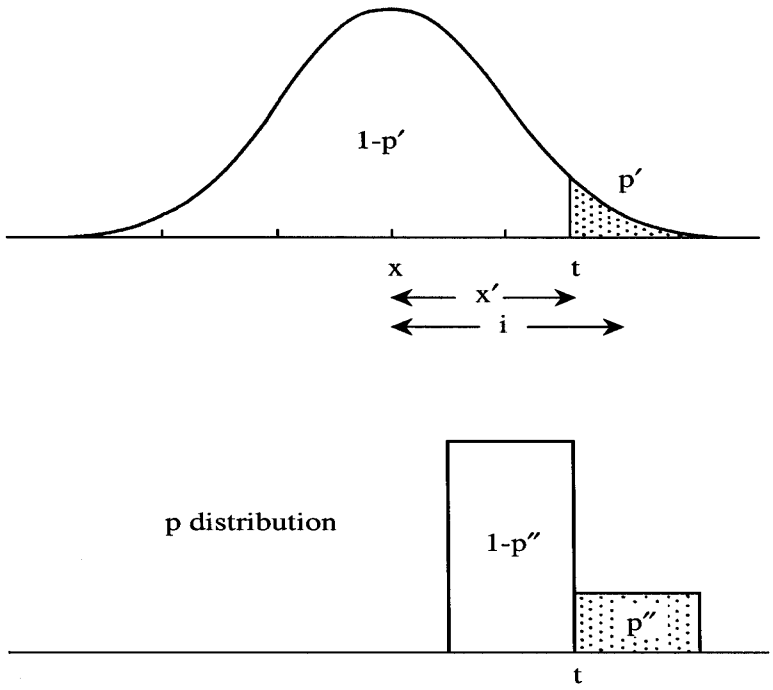


Figure 5.2 Threshold character. Distribution of the underlying variable above and the observable variable below.

Some traits vary in a discontinuous manner. Such traits are only expressed in two, three or a few classes, but are not inherited in a simple mendelian manner. For example mortality and age at maturity show a discontinuous variation. Dempster and Lerner (1950) and Falconer (1965) suggested models that can be used to analyse such data. It is assumed that the observed phenotypic variation on the visible scale (p scale) is determined by an underlying trait or liability (x scale). The expression of the trait on the p scale may be given in two or three classes only. The change on the p scale occurs when the underlying variable surpasses a certain level - the threshold - on the x scale, which is illustrated on Figure 5.2. So the trait, p, shows a discontinuous distribution, but the threshold model assumes that the hypothetical, trait, x is normally distributed. Parameters of the distribution can refer to traits on the visible p scale or on the hypothetical substance of the underlying x scale. The continuous variation of liability is usually both genetic and environmental in origin. In the case of twin births the underlying substance may be a hormone or combination of hormones, which induce polyovulation. In the case of disease resistance the causative substance may be the antibodies resulting from the immune response.

5.3 Variance

5.3.1 ANALYSIS OF VARIANCE

In order to study biological data, the analysis of variance is very useful. By this statistical method it is possible to evaluate the importance of different sources of variation and to test if they contribute significantly to the overall variation observed for a trait.

Table 5.2. Analysis of variance.

Source of variation	DF	Sums of square	Mean square (MS)	E(MS)*
Between classes	n – 1	$\sum d_{\bar{x}}^2$	$\sum d_{\bar{x}}^2 / (n-1)$	$\sigma_e^2 + k\sigma_\alpha^2$
Within classes	n(k - 1)	$\sum\sum d_{x_i}^2$	$\sum\sum d_{x_i}^2 / n(k-1)$	σ_e^2
Total	N - 1	$\sum d_x^2$	$\sum d_x^2 / (N-1)$	

* E(MS) is expectation of mean square.

Analyses of variance for trait X with the balanced design of N observations, n classes and k observations within each class can be calculated as in Table 5.2 assuming the following model:

$$Y_{ij} = \mu + \alpha_i + e_{ij} \tag{5.6}$$

where Yij is an observed variate with $E(\mu) = \mu$ the population mean, α_i is random effects with $\alpha_i = N(0,\sigma_\alpha^2)$ and random error $e \sim N(0,\sigma)$.

The different elements in the analysis of variance can be spelled out as follows:
Total variance, $\hat{\sigma}^2$, which is the square of standard deviation can be calculated as:

$$\hat{\sigma}^2 = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})^2 = \frac{1}{N-1}\left[\sum_{i=1}^{N} x_i^2 - \frac{(\sum_{i=1}^{N} x_i)^2}{N}\right],$$

shortened to:

$$\frac{\Sigma d_x^2}{N-1} \tag{5.7}$$

The statistic:

$$(\Sigma X_i^2 - (\Sigma X_i)^2/N) = \Sigma d^2_X \tag{5.8}$$

is called total sums of square and dividing the sums of square by number of degrees of freedom (N - 1) one gets the mean square or the variance of the trait:

$$(\Sigma X_i^2 - (\Sigma X_i)^2/N)/N\text{-}1 = \Sigma d^2_X/(N-1) \tag{5.9}$$

The mean square between classes for a balanced design ($k_1 = k_2 \ldots = k_n$) is calculated as follows:

$$\hat{\sigma}_{\bar{x}}^2 = \frac{k\sum_{i=1}^{n}(\bar{x}_i - \bar{x})^2}{n-1} = \frac{\frac{1}{k}\sum_{i=1}^{n}(\sum_{i=1}^{k}\bar{x}_i)^2 - \frac{1}{N}(\sum_{i=1}^{N}x)^2}{n-1} = \frac{\sum d_{\bar{x}}^2}{n-1} \tag{5.10}$$

For unbalanced data, the mean square between classes is:

$$\hat{\sigma}_{\bar{x}}^2 = \frac{\sum_{i=1}^{n}k_i(\bar{x}_i - \bar{x})^2}{n-1} = \frac{\sum_{i=1}^{n}\frac{1}{k_i}(\sum_{i=1}^{k}\bar{x}_i)^2 - \frac{1}{N}(\sum_{i=1}^{N}x)^2}{n-1} = \frac{\sum d_{\bar{x}}^2}{n-1} \tag{5.11}$$

Mean squares within classes are calculated as follows for a balanced design:

$$\hat{\sigma}_{x_i}^2 = \frac{\sum_{i=1}^{n}\sum_{i=1}^{k}(x_i - \bar{x}_i)^2}{n(k-1)} = \frac{\sum_{i=1}^{n}[\sum_{i=1}^{k}x_i^2 - \frac{1}{k}(\sum_{i=1}^{k}x_i)^2]^2}{n(k-1)}$$

$$= \frac{\sum d_x^2 - \sum d_{\bar{x}}^2}{n(k-1)} = \frac{\sum_{i=1}^{n}\sum_{i=1}^{k}d_{x_i}^2}{n(k-1)} \tag{5.12}$$

The sums of squares within classes could also be obtained by subtracting the sum of squares between classes from the total sum of squares:

$$\hat{\sigma}_{x_i}^2 = \sum d_x^2 - \sum d_{\bar{x}}^2 \tag{5.13}$$

In analysis of animal breeding data, classes may for example be different sires or different families. A sire variance can then be estimated as:

$$\sigma_\alpha^2 = [\text{MS(between sires)} - \text{MS(within sires)}] / k\,, \tag{5.14}$$

as :
$$E(\text{MS(between sires)}) = \sigma^2 + k\sigma_\alpha^2 \tag{5.15}$$

and:
$$E(\text{MS(within sires)}) = \sigma^2. \tag{5.16}$$

These equations can be expanded to include data from relatives in order to estimate the magnitude of genetic, environmental and maternal effects, see Chapter 9.

5.3.2 VARIANCE OF A SUM

One important elementary algebraic relationship in statistics is that the variance of a sum equals the variance of each component plus two times the covariance between the component parts (Chapman 1962). This rule will be used to study the relationship between the phenotype and its components. It has been shown that the phenotype, P, of an individual consist of two main elements namely its genotype, G, and the environmental components which have influenced the phenotype during lifetime, E. So

$$P = G + E \tag{5.17}$$

The sum of a number of phenotypes equals the sum of their genotypes plus the sum of their environmental components:

$$\Sigma P = \Sigma G + \Sigma E \tag{5.18}$$

the averages of N phenotypes can be calculated:

$$\Sigma P/N = \bar{P} = \Sigma G/N + \Sigma E/N = \bar{G} + \bar{E} \tag{5.19}$$

and the deviation from average can be written in the following way:

$$(P - \bar{P}) = (G - \bar{G}) + (E - \bar{E}) \tag{5.20}$$

The sum of squares is according to the definition:

$$\Sigma(P - \bar{P})^2 = \Sigma[(G - \bar{G}) + (E - \bar{E})]^2 = \Sigma(G - \bar{G})^2 + \Sigma(E - \bar{E})^2 + 2\Sigma(G - \bar{G})(E - \bar{E}) \tag{5.21}$$

And the variance of the sum equals the variance of each component plus two times the covariance:

$$\Sigma(P - \bar{P})^2/(N - 1) = \Sigma(G - \bar{G})^2/(N - 1) + \Sigma(E - \bar{E})^2/(N - 1) + 2\Sigma(G - \bar{G})(E - \bar{E})/(N-1) \tag{5.21}$$

or:
$$\sigma^2_P = \sigma^2_G + \sigma^2_E + 2\mathrm{cov}._{GE} \tag{5.22}$$

and in general:
$$\sigma^2_\Sigma = \Sigma\sigma^2_i + 2\Sigma\mathrm{cov}._{ij} \tag{5.23}$$

If $\mathrm{cov}._{ij} = 0$ then:
$$\sigma^2_\Sigma = \Sigma\sigma^2_i \text{ or } \sigma^2_P = \sigma^2_G + \sigma^2_E \tag{5.24}$$

If $\mathrm{cov}._{ij} = 0$ and $\sigma^2_G = \sigma^2_E = \sigma^2_i$:

$$\sigma^2_\Sigma = n\sigma^2_i \tag{5.25}$$

where n is number of elements in the sum. However, we rarely see that $\sigma^2_G = \sigma^2_E$.

It may, however, be realistic to assume that the covariance between genotype and environment is zero ($cov._{ij} = 0$) when the environments are not too different. A zero correlation means that individuals with different genotypes have equal environmental conditions. Under practical farming conditions with small or moderate environmental differences, there is is often a zero or very low covariance between genotype and environment.

Assuming no interaction between genotype and environment, a phenotypic difference between two individuals with different genotypes remains the same regardless of the type of common environment they have. The absolute phenotypic value for each individual does not have to be the same in the different environments, but the change in phenotype of the individuals are assumed to be the same under different sets of environmental conditions. Sometimes, e.g. when variances are heterogeneous, transformation of data to another scale may eliminate such interactions.

5.3.3 GENETIC VARIANCE

The total genetic value, G, may be divided into different components. The additive genetic value, A, represent the sum of additive genetic effects from each locus and is also called the breeding value. Similarly, dominance genetic value, D, represents the sum of all genetic effects from each locus which is caused by interaction within locus. Epistatic genetic value, I, represent the sum of all interactions between loci. Total genotypic value can be written as the sum of additive, dominance and epistatic genetic effects:

$$G = A + D + I \tag{5.26}$$

Assuming no interaction effects between these components, the total genetic variance will be:

$$\sigma_G^2 = \sigma_A^2 + \sigma_D^2 + \sigma_I^2 \tag{5.27}$$

where σ_G^2 is the variance of total genotype, σ_A^2 is the additive genetic variance, σ_D^2 is the dominance genetic variance and σ_I^2 is the epistatic genetic variance. Estimation of these genetic variances is dealt with in Chapter 9.

5.3.4 ENVIRONMENTAL VARIANCE

A large number of environmental factors affect each animal's traits. The trait may be affected at any stage of development of the individual from fertilisation up until the time that the trait is measured. These environmental factors may be divided in two parts, permanent or systematic factors, $E_{S,}$ and temporary or random factors, $E_{R,}$ as illustrated in Figure 5.4. The random environmental factors may be reduced by repeated measurements of the trait on the same individual or on many individuals within the group of interest (e.g. family). Examples of random environmental factors are random

errors in measurement or recording, random variation in temperature, feeding and management. The effect of each of these factors is very difficult to estimate.

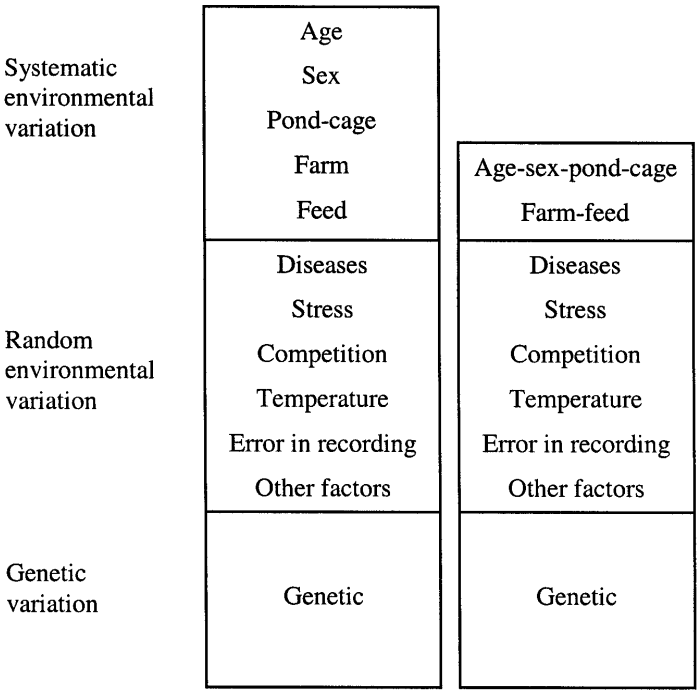


Figure 5.4 Genetic and environmental components of the total variation (left) and after adjustment of systematic environmental factors (right) of a trait.

Systematic environmental effects cannot be reduced by repeated measurements and represent effects, which are common to certain groups of animals or effects that vary according to year. Examples of systematic effects are effects of age, sex, temperature, pond, cage, farm, and feed. In breeding experiments or breeding programs, it is important that we strive for testing the animals under equal environmental conditions, attempting to minimise random and systematic environmental effects. This will ensure that the animal's phenotype is a better predictor of its genotype.

However, it is not possible to avoid all environmental effects. In order to reduce the influence of environmental factors, the records should be corrected for systematic environmental effects. Data sets for estimation of correction factors should be large and should have a certain number of records in each group to assure a low s.e. of the estimates. If for example the effect of sex on body weight is to be corrected for, mean weight for each sex can be used as correction factor if it is not confounded with other effects. If the effects are confounded, a statistical model including all systematic fixed

and random effects should be used to estimate the effects. Measurements from different groups may be adjusted to a common mean or to the group closest to the population mean. An example of the derivation of correction factors is given below using the data in Table 5.3.

Table 5.3 Mean ($\bar{X}$) standard deviation (σ) and coefficient of variation (CV) for body weight of male (1), female (2) and none mature fish (3)

Cage	Sex	$\bar{X}$ (Rel.weight)	σ	CV
Salmon				
1	1	4.75 (100)	1.09	23
	2	5.80 (122)	1.28	22
	3	2.92 (61)	0.97	33
2	1	4.76 (100)	1.08	23
	2	5.86 (123)	1.27	22
	3	3.15 (66)	0.97	31
Rainbow trout				
1	1	3.71 (100)	0.85	23
	2	4.06 (109)	0.92	23
	3	3.09 (83)	0.95	31
2	1	3.59 (100)	0.86	24
	2	3.96 (110)	0.93	23
	3	3.04 (85)	0.97	32

For salmon in cage 1 when adjusting to the females' body weight they will be as follows:

Correction factor for males: 4.75 - 5.80 = - 1.05
Correction factor for non-mature fish: 4.75 - 2.92 = + 1.83

When applying this additive correction, we add 1.83 kg to the body weight of all non-mature fish and subtract 1.05 kg from the body weight of all males. The result of the additive correction is that all groups will have the same average. However, the additive correction has no effect on the variance within groups. The variation usually increases at higher mean values. This scale effect can be seen in Table 5.3 where body weight for the different sex groups has different standard deviations. Because males have the largest standard deviation, the males should also have the greatest range of adjusted body weights after additive correction. Estimates of breeding values based on such data and the ranking of potential breeders will be biased because of such heterogeneous variance between groups. In order to solve this problem we can use multiplicative correction. It implies that observations are corrected by multiplying the observation with a factor

above or below one. Such correction factors affect the variance of the observations. Correction factors above one increase the variance, whereas factors below one reduce it. If two groups have similar coefficients of variation (which is true for males and females in Table 5.3) the averages as well as standard deviations will be similar after the multiplicative correction has been made. This implies that body weight of males must be divided by a factor greater than one. Here, the factor will be 1.22 (5.80/4.75). For non-mature fish the situation is different, since the standard deviation for males and non-mature fish are quite similar (however the coefficient of variation is different because the mean of the male fish is larger). When observations in different classes have different variance not due to the scale effect shown above, we can correct by multiplying the observations in class i with a factor:

$$\sigma_i = \sigma_o / \ \sigma_i \tag{5.28}$$

where σ_i = standard deviation estimated for class i and σ_o = desired standard deviation.

When the systematic environmental effect is a continuous variable, regression analysis should be used to estimate the correction factors. Variation in age in days from first feeding to recording is an example. The coefficient of regression of body weight on age should be used to estimate the correction factor. If the regression is curvilinear, multiple regressions should be estimated and used for adjustment.

Corrected values can also be obtained by using the observations' deviation from appropriate class means. The class means for example for a pond is then estimated using the same data as the data which should be corrected. This method is appropriate for effects on big groups of individuals and effects that can not be obtained from accurate historical data. Effects of pond, year and season are examples of such effects.

Data should be adjusted before estimation of breeding values, phenotypic and genetic parameters. The effect of correcting records for systematic environmental factors on the variance is shown in Figure 5.4. When using BLUP methodology for prediction of breeding values, as described in chapter 16, correction of systematic environmental effects is carried out simultaneously with prediction of breeding values. This ensures unbiased prediction of breeding values as well as unbiased estimates of correction factors. Also, heterogenous variance can be accounted for in the BLUP procedure.

5.3.5 MATERNAL VARIANCE

Maternal ability is a source of variance from dam effects, which can be both environmental as well as genetic in origin. In mammals, we find both prenatal and postnatal maternal effects, and in this case maternal influences can be substantial. In fish and shellfish, however, maternal effects are mostly related to variation in egg size and egg quality resulting in variation in hatching and survival during the first stages of growth and development. These effects may be caused by deficiency in nutrition of dams, especially in vitamins, or water quality, water temperature and management practises.

Egg size has been found to affect growth rate in fish during the first period of life, in common carp (Kirpichnikov, 1970), Atlantic salmon (Nævdal et al., 1975; Refstie personal com.), rainbow trout (Gall, 1972; 1974), tilapia (Siraj et al., 1983) and channel catfish (Reagan and Conley, 1977). Refstie (personal comm.) did not find any effect of egg size on mortality of fry and fingerlings in salmon. However, there are examples of broodstock fed low levels of vitamin C with increased mortality in fry. Some shellfish like crayfish carry their eggs under abdomen until hatching and thus the female will influence the larvae for a longer period of time than l in fish. The mouth brooding in tilapia may also result in larger maternal effect compared to none mouth-brooders. However, maternal effects have not been shown to be of significant importance in fish and shellfish.

In a breeding program where families are reared separately until tagging, environmental differences between families may occur. Such environmental differences are also called tank or hapa effects depending on the technology used, and may be a combination of maternal and environmental effects. An estimate of 5 % of the total variation in body weight of Atlantic salmon fingerlings was found at an age of 190 days (Refstie and Steine, 1978). However, the tank effect was reduced to less than one percentage after the fish had been reared for two years in the sea (Gunnes and Gjedrem, 1978). Higher estimates (7 to 13 %) have been reported for tilapia (Bentsen personal communication).

Table 5.4. Analysis of variance including interaction.

Source of variation	DF	Sums of square	Mean square
Between strains	r - 1	$\sum d^2_{\bar{x}_i}$	$\sum d^2_{\bar{x}_i} /(r-1)$
Between environments	m - 1	$\sum d^2_{\bar{x}_j}$	$\sum d^2_{\bar{x}_j} /(m-1)$
Interaction strain – environ.	(r-1)(m-1)	$\sum d^2_{\bar{x}_{ij}}$	$\sum d^2_{\bar{x}_{ij}} /(r-1)(m-1)$
Within classes	rm(n - 1)	$\sum\sum d^2_{x_{ij}}$	$\sum\sum d^2_{x_{ij}} /rm(k-1)$
Total	N - 1	$\sum d^2_{x_{ijk}}$	$\sum d^2_{x_{ijk}} /N-1$

5.3.6 GENOTYPE-ENVIRONMENT INTERACTION

A common practise in research experiments is to raise different strains of fish in different environments. Such data provides an opportunity to study genotype-environment interaction. The analysis of variance may be performed as shown in Table 5.4.

Interaction between strain and environment can also be called genotype-environment interaction since strains will belong to different genetic groups. As mentioned above, the meaning of a significant genotype-environment interaction is that the difference between strains or genetic groups is not the same in different environments. However, the interaction will only have practical consequences when the difference is so big that it leads to reranking of the genetic groups in another environment. In Figure 14.4, a low genotype-environment interaction is demonstrated (Eknath et al.,1993). Different strains of tilapia were tested in 11 different environments. In the figure, the environments with lowest growth rate are to the left and those with the highest growth rate to the right. Body weight of each strain is presented as the deviation from the average in each environment. As can be seen the ranking of strains are very similar in the different environments, especially in the good environments. Eknath et al. (1993) estimated the G-E interaction to account for 0.3 % of the total variation. An example of significant genotype-environment interaction is shown for the extreme environments leading to reranking in Figure 14.1. Comparing the less extreme environments, the difference between the genotypes is smaller and the ranking of the genotypes is the same.

There are varying results concerning the importance of genotype-environment interaction in fish and shellfish, and it is difficult to draw a general conclusion about the number of breeds or populations to be used in a breeding program. Evidence for G - E interaction should therefore be investigated for each individual breeding program.

5.4 Relationship between traits

5.4.1 CORRELATION AND REGRESSION

Relationship between two traits, X and Y, may be expressed as covariance (cov_{XY}), correlation (r_{XY}) or regression (b_{XY}):

The covariance is: $$cov_{XY} = \Sigma(X - \bar{X})(Y - \bar{Y})/N - 1) \quad (5.29)$$

Where X and Y are observations on two traits. The correlation between the two is $r_{XY} = cov_{XY}/\sigma_Y\sigma_X$, which is the covariance divided by the product of the standard deviations of the traits or geometric mean of variances:

$$r_{XY} = [\Sigma(X - \bar{X})(Y - \bar{Y})/N - 1]/\sqrt{(\Sigma d^2_X)/(N - 1)(\Sigma d^2_Y)/(N - 1)}$$
$$= \Sigma(X - \bar{X})(Y - \bar{Y})/\sqrt{(\Sigma d^2_X)(\Sigma d^2_Y)} = (\Sigma XY - \Sigma X\Sigma Y/N)/\sqrt{(\Sigma d^2_X)(\Sigma d^2_Y)} \quad (5.30)$$

The coefficient of regression of Y on X is:

$$b_{YX} = cov_{XY}/\sigma_X^2 = r_{YX}\sigma_Y/\sigma_X = \Sigma d_{YX}/\Sigma d_x^2 \quad (5.31)$$

A simple regression equation:

$$Y = a + b_{YX} X \quad (5.32)$$

describes a straight line with b_{YX} being the slope of the line and a representing the value of $\overline{Y}$ when $X = 0$. The regression line passes through the means of both traits, such that the estimate of the intercept, a, is:

$$a = \overline{Y} - b_{yx}\overline{X} \tag{5.33}$$

Hence, the regression equation can be written as:

$$Y = (\overline{Y} - b_{YX}\overline{X}) + b_{YX}X \tag{5.34}$$

or as:

$$Y - \overline{Y} = b_{YX}(X - \overline{X}) \tag{5.35}$$

The estimated regression coefficient is such that the sum of the squared deviations between the observed values of Y and the regression line are minimised. For further information see Snedecor and Cochran (1980). In animal breeding, we can for instance use the estimated regression equation to predict an animal's breeding value from its phenotype.

5.4.2 ANALYSIS OF COVARIANCE

The between class correlation coefficient, for example the between sire or family correlation, can be obtained from an analysis of covariances. It is equal to the sum of cross products between classes divided by the geometric mean of sum of squares between classes:

$$r = (\Sigma(\Sigma Y_i \Sigma X_i)/k_i - \Sigma Y \Sigma X/N)\sqrt{\Sigma d^2_Y} \,.\Sigma\, d^2_X = \Sigma d_{YX}/\sqrt{\Sigma d^2_X}\sqrt{\Sigma d^2_Y} \tag{5.36}$$

The within class correlation coefficient is:

$$r_{YiXi} = \Sigma d_{YX} - \Sigma d_{YX}/\sqrt{\Sigma d^2_{Yi}}\,\Sigma d^2_{Xi} = \Sigma d_{YiXi}/\sqrt{\Sigma d^2_{Xi}} \tag{5.37}$$

The regression sum of squares is the sum of squares explained by the regression, and can be expressed as a function of the correlation coefficient:

$$\Sigma\, d^2_Y \Sigma\, d^2_X r^2_{YX} = [\Sigma\,(Y - \overline{Y})(X - \overline{X})]^2 = (\Sigma\, d_{YX})^2 \tag{5.38}$$

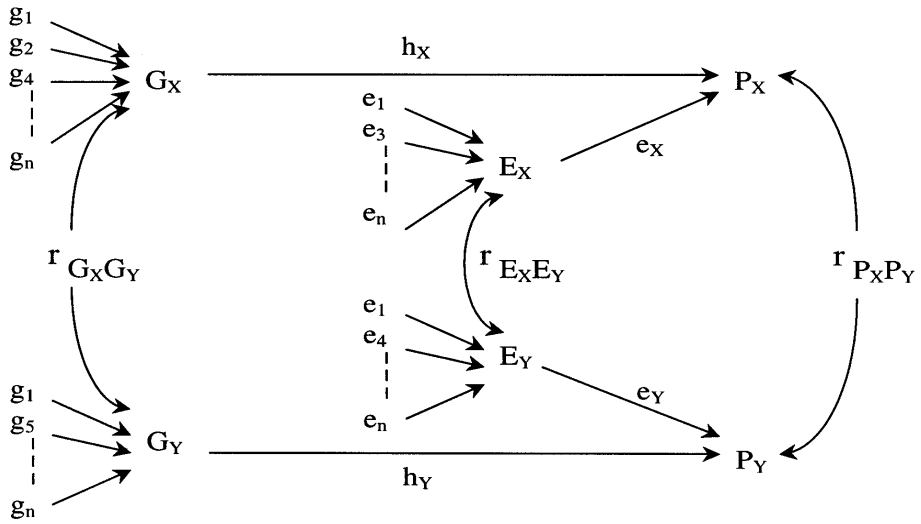


Figure 5.5 Relationship between genotypes (G_x and G_y), environments (E_x and E_y) and phenotypes (P_x and P_y). $G_{1\text{-}gn}$ are single genes and $e_{1\text{-}en}$ are environmental factors. R_{PxPy} are phenotypic correlation, r_{ExEy} is environmental correlation and r_{GxGy} is genetic correlation. h_x is path from genotype to phenotype for trait P_x or the square root of the heritability (h_x^2). Ex is path from environment to the phenotype for trait P_x or the square root of the environmental variance (e_x^2).

The sums of squares independent of the regression or sums of squares of prediction errors is:

$$\sum d_Y^2 (1 - r_{YX}^2) = \sum (Y - \hat{Y})^2 \tag{5.39}$$

where

$$\hat{Y} = \hat{a} + \hat{b}_{YX} X , \tag{5.40}$$

the value fitted or predicted by the estimated regression equation. The standard error of the estimate of the regression coefficient is:

$$se(b_{YX}) = \sqrt{\frac{\sum d_Y^2 (1 - r_{YX}^2)}{(N-2)\sum d_X^2}} \tag{5.41}$$

5.4.3 GENETIC AND PHENOTYPIC CORRELATION

An observed phenotypic correlation may be due to environmental and (or) genetic covariance. The covariance between two traits can be estimated at both the phenotypic and genetic levels, in a manner similar to estimation of phenotypic and genetic variance of one trait. Genetic covariance and standard deviations are combined to estimate the genetic correlation:

$$r_{G(XY)} = \frac{\sigma_{G(XY)}}{\sigma_{G(X)}\sigma_{G(Y)}} \tag{5.42}$$

Assuming that the phenotype is the sum of additive genetic and environmental effects, the phenotypic covariance ($\sigma_{P(XY)}$) between two traits (X and Y) is the sum of the additive ($\sigma_{G(XY)}$) and environmental covariances ($\sigma_{E(XY)}$):

$$\sigma_{P(XY)} = \sigma_{G(XY)} + \sigma_{E(XY)} \tag{5.43}$$

The main causes of genetic correlation are pleiotrophy and linkage. A gene usually has its main effect on one trait, but frequently genes affect several traits, and this relationship is called pleiotrophy. The gene effect may also change during the life, e.g. be positive in one stage of life and negative in another. When two or more genes are located closely on the chromosome and linked together, they may segregate as one unit and thus act as one gene in several generations, again causing a genetic correlation.

Genetic correlations vary between -1 and 1 (as do all correlations). Genetic correlation's equal or close to one indicates that the two traits are genetically very similar. The genetic correlation between two genetically independent traits is zero.

As an example of a high genetic correlation, the genetic correlation between body weight (W) and body length (L) in rainbow trout is estimated to be 0.98. (Gunnes and Gjedrem, 1978). The environmental correlation was estimated to 0.86 and the phenotypic correlation to 0.88 using the same data. These relationships are shown in Figure 5.6.

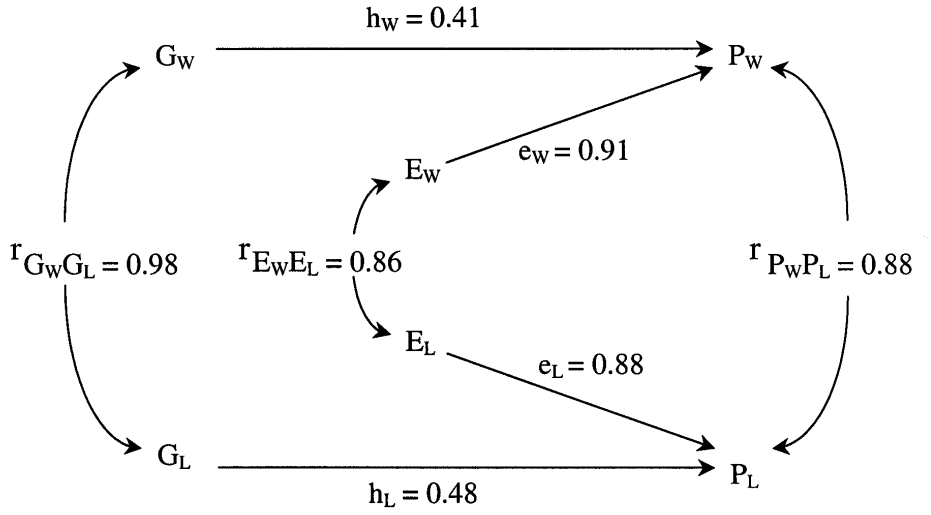


Figure 5.6 Relationships between body weight (W) and body length (L) in rainbow trout.

Some estimates of genetic and phenotypic correlations are given in Table 5.5 and Table 5.6. The phenotypic, genetic and environmental correlations and ratios of variances are indicated in Figure 5.5. The phenotypic correlation can be calculated as follows:

$$r_{P(XY)} = h_X \; h_Y \; r_{G(XY)} + e_{1X} \; e_Y . \; r_{E(XY)} \qquad (5.44)$$

where $r_{E(XY)}$ = environmental correlation and e_X and e_Y are the square roots of the

Table 5.5 Phenotypic (below diagonal) and genetic (above diagonal) correlations between body weight and quality traits.

Trait	Body weight	Fat %	Area of fat depot	Flesh colour	Condition factor	Dressing %	Texture
Body weight		0.42 (0.17) [2] -0.19 [3] 0.40 (0.35) [4] 0.65 (0.33) [5] 0.22 [7] -12 (0.26)[9]	0.17[8]	-0.21 (0.33) [1] 0.31 (0.20) [2] 0.21 [3] 0.58 [7] 0.36 (0.21)[9]	0.37 (0.18) [2] 0.16 [3] -0.38 (0.18) [5] 0.29 (0.17)[9]	0.61 (0.17) [2] 0.07 [3] 0.04 (0.18)[9]	0.15[10]
Fat %	0.63 [2] 0.21 [3] 0.45 (0.05) [4] 0.36 (0.05) [5] 0.67[6] 0.30 [7]		0.71[8]	-0.39 (0.32) [2] -0.44 [3] 0.22 (0.60)[5] -0.03 [7] 0.13(0.24)[9]	0.64 (0.14) [2] 0.15 [3] 0.05 (0.80) [4] 0.00 (0.20)[9]	0.59 (0.19) [2] 0.24 [3] 0.38 (0.18)[9]	0.19[10]
Area of fat depot	0.53 [6]	0.86 [6]					0.10[10]
Flesh colour	0.19 (0.02) [1] 0.17 [2] 0.07 [3] 0.64 [7] 0.13[9]	0.03 [2] -0.23 [3] -0.05 (0.05) [5] 0.33 [7] 0.11[9]			-0.25 (0.27) [2] -0.08 [3] 0.34 (0.17)[9]	-0.39 (0.32) [2] -0.45 [3] 0.00 (0.18)[9]	0.14[10]
Condition factor	0.51 [2] 0.27 [3] 0.50(0.05) [5] 0.42[9]	0.45 [2] 0.16 [3] 0.09 (0.05) [4] 0.38[9]		0.04 [2] 0.15[9]		0.51 (0.19) [2] -0.09 [3] -0.27 (0.14)[9]	
Dressing %	0.26 [2] 0.29 [3] -0.20[9]	0.09 [2] 0.13 [3] 0.07[9]		0.05 [2] -0.22 [3] 0.13[9]	0.12 [2] -0.02 [3] -0.22[9]		

1) Rye et al. (1994), Atlantic salmon
2) Rye and Gjerde (1996) Atlantic salmon
3) Gjerde and Schaeffer (1989) Rainbow trout
4) Elvingson (1992) Rainbow trout
5) Elvingson and Nilsson (1992) Arctic char
6) Rye pers. med. Atlantic salmon
7) Iwamoto et al. (1990) Coho salmon
8) Gjerde pers. med. Atlantic salmon
9) Kause et al. (2002) Rainbow trout
10) Refstie et al. (1999) Atlantic salmon

environmental portion of the phenotypic variance, $\sqrt{(1-h^2)}$ for X and Y, respectively. The phenotypic correlation, $r_{P(XY),}$ consists of a genetic part, h_X h_Y $r_{G(XY),}$ and an environmental part, e_X e_Y. $r_{E(XY),.}$ The path from genotype to phenotype (h_i) is the square root of the heritability. In order to predict the response to selection we need these parameters. It is particularly important to know the magnitude of genetic correlations when predicting breeding values.

The genetic part of the phenotypic correlation contains the additive genetic variance and covariance. The estimates of the environmental portion may also include most of the epistatic and dominance variance and covariance.

Table 5.6 Genetic correlations between growth rate and different measurements of survival.

	Genetic correlation	Spesies	Number	
			Sire	Dam
1)Overall survival, fingerlings	0,30	B.t.	32	32
2)Coldwater vibiosis, adult	0,18	A.s.	53	329
3)Overall survival, fingerlings	0,37	A.s.	187	1404
3)Overall survival, fingerlings	0,23	Rb.	213	1062
4)Furunculose, fingerlings,	0,30	A.s.	25	50
challenge	0,31	A.s.	100	298
5)Overall survival, fingerlings	0.50	A.c.	36	32
6)Fungal infection	- 0.12	P.v.	57[D)]	279[F)]
7)Taura Syndrom Virus	0.20	N.t.	50	150
8)Overall survival, 90 days				

A.s. = Atlantic salmon
Rb = Rainbow trout
B.t. = Brook trout
A.c. = Arctic char
P.v. = Penaeus vannamei
N.t. = Nile tilapia
D) Dams, maternal half-sibs
F) Full-sibs

1)Robison & Luempert (1984)
2) Standal & Gjerde (1987)
3)Rye et.al. (1990)
4)Gjedrem et.al. (1991)
5)Jonasson (1993)
6)Nilsson (1992)
7)Fjalestad et al. (1997)
8)Eknath et al. (1998)

5.5 Heritability

Heritability is one of the most useful parameters in animal breeding. The heritability indicates the proportion of the total phenotypic variation which is genetic in origin. It is important to know the size of the heritability when planning a breeding program as well as when predicting the response to selection or individuals' breeding values. The heritability can vary between 0 and 1. Heritability in the broad sense, h_B^2, is defined as the ratio of the total genetic variance, σ_G^2, to the total phenotypic variance, σ_P^2:

$$h_B^2 = \sigma_G^2/\sigma_P^2 \qquad (5.45)$$

However, since σ_G^2 also may contain dominance and epistatic effects which do not respond to selection, heritability in the broad sense is not often used. Heritability in the narrow sense is defined as the ratio of the additive genetic variance, σ_A^2, and the phenotypic variance, σ_P^2:

$$h^2_A = \sigma_A^2/\sigma_P^2 \tag{5.46}$$

Thus the narrow sense heritability is the portion of the total variance that is due to differences between the breeding values of individuals in the population. A high heritability indicates that a substantial portion of the phenotypic variance is due to additive genetic variance, and that a larger selection response can be expected. When the heritability is low, environmental effects, make a larger contribution to the phenotypic variance than additive genetic effects.

It should be mentioned that the heritability is not a general and static characteristic of a breeding population. It is only relevant to the data used for estimation, or the population from which the data was sampled. When environmental conditions are changed and or if a population responds to selection (with a corresponding change in the additive variance), the heritability may change. Some estimates of heritability for different species and different traits are given in Table 5.7. Different methods for estimation of the heritability and other phenotypic and genetic parameters are presented and discussed in Chapter 9.

5.6 Path coefficients

The path coefficient procedures were developed by Wright (1934). They can be used in connection with problems in which we have a logical point of view in regard to which variables may be looked on as cause and which as effect. The definition given by Wright for a path coefficient was: «Path coefficient, measuring the importance of a given path of influence from cause to effect, is the ratio between the standard deviation of the effect when all causes are constant except the one in question, the variability of which is kept unchanged, to the total standard deviation». The path coefficient squared has been called the coefficient of determination since it is a measure of the percentage of variation in one variable which is accounted for by variation in another. Similarly the square of a simple coefficient of correlation is called, coefficient of determination.

The path diagram is constructed as follows:

- A streight arrow is drawn to each dependent variable from each of its sources.
- A curved, double-headed arrow is drawn between each pair of independent variables thought to have non-zero correlation.

In Figure 5.5; 5.6; 5.7 and 5.8 the use of path diagram is illustrated.

Table 5.7 Phenotypic means (X), coefficient of variations (CV), heritabilities (h^2) with standard error (se), estimated from sire (S), dam (D) and full-sib (F) component of variance for flesh quality traits in fish.

Species/traits	X	CV	h^2_S (se)	h^2_D (se)	Number of families	Authors
Rainbow trout:						
Mature weight, kg	1.9	25	0.20 (0.10)	0.57	49_S; 192_D	Gall and Huang (1988a)
2½ years body weight, kg	3.4	21	0.21	0.41	47_S; 249_D	Gjerde and Schaeffer (1989)
Dressing %	83.3	6	0.01 (0.05)	0.14 (0.07)	13_S; 108_D	Gjerde and Gjedrem (1984)
Carcass quality, score	3.6	20	0.14 (0.06)	0.07 (0.05)	56_S; 108_D	”
Flesh colour, score	3.4	23	0.06 (0.08)	0.28 (0.09)	56_S; 108_D	”
Dressing %	87.3	2	0.36	0.42	47_S; 249_D	Gjerde and Schaeffer (1989)
Condition factor	1.6	10	0.19	0.34	47_S; 249_D	”
Flesh colour, score	4.3	15	0.27	0.29	47_S; 249_D	”
Fat %	14.8	17	0.47	0.19	37_S; 111_D	”
Protein %	20.0	7	0.03	0.08	37_S; 111_D	”
Condition factor	1.7	11	0,03 (0.04)		32_S; 81_D	Elvingson (1992)
Fat %	14.7	23	0.17 (0.09)		32_S; 81_D	”
Survival, unspecifies			0.16 (0.03)		186_S; 770_D	Rye et al. (1990)
VHS			0.69 (0.25)		14_S	Chevassus and Dorson (1990)
Age at maturity			0.21 (0.14)		14_S; 14_D	McKay et al. (1986)
Age at sexual maturity			0.13		56_S	Gjerde and Gjedrem (1984)
Age at sexual maturity			0.07		98_S	Gjerde (1986)
Age of sexual maturity			0.12-0.34 (0.03)		340_S; 552_D	Kause et al. (2003)
Body weight, g			0.21-0.27(0.03)		340_S; 552_D	”
Growth	1.6	28	0.26 (0.12)		44_S; 145_D	Kinghorn (1983b)
Food consumption	1.4	24	0.41 (o.13)		44_S; 145_D	”
Gross food conversion	1.2	8	0.03 (0.10)		44_S; 145_D	”
Fat %	9.7	10	0.47 (0.34)		44_S; 145_D	”

Table 5.7 continues:

Atlantic salmon*:*						
Body weight, kg	6.6	29	0.35 (0.10)	0.32 (0.09)	58_S; 171_D	Rye and Refstie (1995)
Dressing %	90	4	0.03 (0.02)	0.02 (0.02)	105_S; 248_D	Gjerde and Gjedrem (1984)
Carcass quality, score	3.8	19	0.16 (0.05)	0.14 (0.03)	105_S; 248_D	”
Flesh colour, score	3.6	16	0.01 (0.03)	0.17 (0.04)	105_S; 248_D	”
Fat %	15.6	16	0.30 (0.09)	0.38 (0.10)	58_S; 171_D	Rye and Gjerde (1996)
Dressing %	93.9	1	0.20 (0.07)	0.18 (0:08)	58_S; 171_D	”
Flesh colour, score	3.3	18	0.09 (0.05)	0.16 (0.08)	58_S; 171_D	”
Condition factor	1.21	13	0.31 (0.09)	0.30 (0.09)	58_S; 171_D	”
Flesh color	7.7	18	0.47 (0.13)			Rye et al. (1994)
Fat %	22.0	17	0.46 (0.21)	0.43 (0.17)	30_S; 81_D	Rye pers. med.
Area of fat depot	8.3	21	0.41 (0.19)	0.27 (0.15)	30_S; 81_D	”
Area of fat depot			0.32 (0.15)			Gjerde pers.med.
Survival, unspecified			0.08 (0.02)		178_S; 1127_D	Rye et al. (1990)
Vibriose			0.12 (0.05)		42_S; 140_D	Gjedrem and Aulstad (1974)
Coldwater vibriose			0.13 (0.08)		32_S; 81_D	Gjedrem and Gjøen (1995)
Furunculosis			0.48 (0.17)		25_S; 50_D	Gjedrem et al (1991)
Furunculosis			0.16 (0.12)		32_S; 81_D	Gjedrem and Gjøen (1995)
BKD			0.23 (0.10)		32_S; 81_D	Gjedrem and Gjøen (1995)
Sea lice, *Lepeophtheirus*			0.25 (0.07)		25_S; 50_D	Kolstad et al. (2004)
Sea lice, *Caligus*			0.22		73_F	Mustafa and MacKinnon (1999)
Age of sexual maturity			0.48 (0.20)		Realized	Gjerde (1984a)
Age at sexual maturity			0.20		84_S	Gjerde and Gjedrem (1984)
Age of sexual maturity			0.07		50_S	Gjerde (1986)

Table 5.7 continues:

Coho:						
Body weight, g	239	41	0.19 (0.11)		20_S; 40_D	Hershberger et al. (1990)
Body weight, g	371	40	0.33 (0.10)		20_S; 40_D	"
Colour	418	36		0.50 (0.16)	35_F	Iwamoto et al. (1990)
Colour	464	28		0.30 (0.14)	38_F	"
Fat %	8	63		0.18 (0.13)	35_F	"
Fat %	9	52		0.19 (0.23)	38_F	Cerda et al. (2003)
Body weight			0.36 (0.03)		120_S	"
Colour			0.11 (0.02)		120_S	"
Fat thickness			0.04 (0.03)		120_S	Beacham and Evelyn (1992)
Furunculosis			0.00 (0.12)		14_S; 28_D	
Chinook:						
Body weight, kg	2	26	0.25 (0.10)		16_S, 32_D	Winkelman and Peterson (1994)
Furunculosis			0.14 (0.11)		15_S, 30_D	Beacham and Evelyn (1992)
Common carp:						
Body weight			0.25			Smisek (1979)
Rohu carp:						
Body weight, g	440	31	0.23 (0.06)		311_S; 357_D	Gjerde et al. (2004)
Survival			0.16		57_S; 20_D	"
Catfish:						
48 week weight, female, g			0.52 (042)		13_S; 26_D	El-Ibiary and Joyce (1978)
48 week weight, male, g	237	24	0.27 (0.37)	0.51 (0.37)	13_S; 26_D	"
Dressing %, female	328	27	0.00 (0.20)	0.67 (0.41)	13_S; 26_D	"
Fat %	69	2	0.08 (0.25)	0.47 (0.36)	13_S; 26_D	"
Dressing %, male	42	9	0.00 (0.27)	0.50 (0.32)	13_S; 26_D	"
Fat %	68	2	0.00 (0.23)	0.84 (0.44)	13_S; 26_D	"
Fat %	43	7	0.61 (0.78)	0.64 (0.37)	10_S; 20_D	Reagan (1980)
Feed intake				0.41	31_F	Silverstein et al. (2001)

Table 5.7 continues:

Tilapia:						
90 day weight, g	19	27	0.04 (0.06)	0.04 (0.08)	16_S; 32_D	Tave and Smitherman (1980)
Body weight			0.32		Realized	Jarimopas (1990)
16 weeks, weight, g			0.38			Bolivar (1999)
90 day weight, g			0.23	0.53	50_S; 150_D	Eknath et al. (1998)
Survival, %			0.08		50_S; 150_D	”
Body weight, cages	181		0.30	0.83	52_S; 84_D	Pondzoni et al. (2003)
Body weight, ponds	291		0.39	0.93	52_S; 84_D	”
Body weight, 98-day			0.20 (0.04)			Gall and Bakar (2002)
Arctic char*:*						
Body weight, g	651	36	0.52 (0.17)	0.38 (0.14)	36_S; 32_D	Nilsson (1992)
Body weight, g	439	37	0.45 (0.19)	0.45 (0.20)	36_S; 32_D	"
Condition factor	2	11	0.24 (0.15)		12_S; 21_D	Elvingson and Nilsson (1992)
Fat %	20	12	0.06 (0.08)		12_S; 21_D	“
Colour	1	46	0.28 (0.17)		12_S; 21_D	”
Fungi			0.34 (0.14)		36_S; 32_D	Nilsson (1992)
Turbot:						
Body weight, g	829	36	0.70 (0.19)	0.45 (0.28)	35_S; 19_D	Gjerde et at. (1997)
Atlantic cod:						
Body weight, g	25	38		0.29 (0.27)	51_F	Gjerde et al. (2004)
Survival, %				0.00 (0.14)	51_F	”
Shrimp:						
Body weight, g	4	70	0.10 (0.002)	0.39 (0.004)	17_S; 34_D	Benzie et al. (1997)
Body weight, g	18	17	0.76 (0.15)	0.71 (0.05)	57_D; 279_S	Fjalestad et al. (1997)
Taura Syndrom Virus	48	34	0.22 (0.09)	0.35 (0.03)	57_D; 279_S	"
Body weight, g	16	22	0.40 (0.09)		26s; 52D	Suarez et al. (1999)
Survival, %	66	-	0.13 (0.05)		26s; 52D	”

Table 5.7 continues:

Oysters:						
Body weight, g	24	34	0.28 (0.06)		Realized	Jarayabhand et al. (1995)
Shell length, mm			0.34 (0.12)			Toro and Newkirk (1991)
Shell length, mm			0.19 (0.07)			Toro and Newkirk (1990)
Live weight			0.11-0.24			"
Body weight, g			0.33 (0.19)			Lannan (1972)
Body weight			0.08 (0.25)			Haley et al. (1975)
Yield per bag			0.54			Langdon et al. (2000)
Body weight			0.10 – 0.51		25_F	Davis (2000)
Scallop:						
Body weight			0.37		Realized	Heffernan et al. (1993)
Body weight, g			0.43 (0.09)			Ibarra et al. (1999)
Body weight, g	47	26	0.59 (0.13)			"
Body weight, g	61	21	0.21			Crenshaw et al. (1991)
Clam:						
Body size, mm	33	17	0.42		Realized	Hadley et al. (1991)
Growth			0.40		Realized	Chrenshaw et al. (1996)
Abalone:						
Shell length, mm	51		0.34		29_S; 88_D	Jonasson et al. (1999)
Survival, %	10		0.11		29_S; 88_D	"
Mussel:						
Body length, mm	30	8	0.09		20_F	Brichette (2001)
Shell length, mm	30	20	0.22 (0.07)		20_S; 60_D	Mallet et al. (1986)
Shell length, mm	24	31	0.92 (0.27)		20_S; 60_D	"

If we consider three traits, then one possible path coefficient is a partial regression coefficient of a trait A on a trait B while keeping another trait C constant, Figure 5.7. The correlation between trait A and B is equal to the path of A on B when the correlation between trait B and C are zero.

Similarly, a path coefficient can be described for the relationship between a parent and a progeny by seperating this coefficient into two parts:

1) The path from a gamete to the zygote, a, which the gamete forms when it unites with another gamete and
2) The path from the parental zygote to the gamete it produces, b.

The genetic value of the offspring is the sum of the genetic values of the two gametes (X and Y) which unite in the offspring, so the path coefficient a_1 and a_2 in Figure 5.6 is:

a_1 = the path from the sperm to the offspring $= \sigma_x/\sigma_{x+y} = \sqrt{\sigma^2_x}/(\sigma^2_x + \sigma^2_y + 2\text{cov}_{xy})$

a_2 = the path from the ovum to the offspring = σ_y/σ_{x+y}.

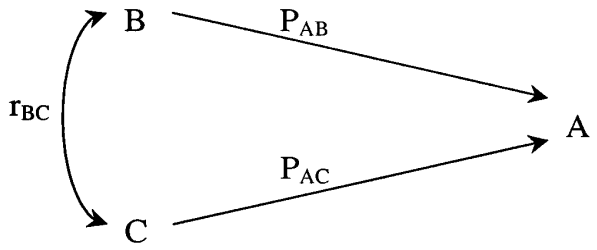


Figure 5.7 A path diagram with directional arrow from B and C to A and bi-directional arrow between B and C.

The ova and sperm are likely to be equally variable and if σ_x and σ_y are similar we can drop the subscripts. If the parents (ova and sperm) are related and the offspring (zygote) have an inbreeding coefficient, F, the path from the ovum (sperm) to the zygote will be:

$$a = \sqrt{\sigma^2_x}/(2\sigma^2_x + 2F\sigma^2_x) = \sqrt{1}/2(1 + F) \tag{5.47}$$

And b = path from parental zygote (with gametes X and W) to the gamete it produces (X) so the path from the sire to it's gamete is:

$$b_S = r_{SX} = \text{cov } X(X + W)/\sigma_x\sigma_S = (\sigma^2_x + \sigma_x\sigma_W F_S)/\sigma_x\sigma_S = (\sigma_x + \sigma_W F_S)/\sigma_S \tag{5.48}$$

If $\sigma_S = \sigma_W = \sigma_x$, the formula for b will reduce to:

$$b = (1 + F_S)/\sqrt{2}(1 + F_S) = \sqrt{(1 + F_S)/2} \tag{5.49}$$

The direct path from parent to offspring is the product of the path from parental zygote to gamete and the path from gamete to offspring zygote as illustrated in Figure 5.8. The path from sire to offspring will be:

$$a_1 b_S = ½\sqrt{(1 + F_S)/(1 + F)} \quad (5.50)$$

and from dam to offspring: $$a_2 b_D = ½\sqrt{(1 + F_D)/(1 + F)} \quad (5.51)$$

If sire and dam are equally inbred, the equation will be:

$$ab = ½\sqrt{(1 + F')/(1 + F)} \quad (5.52)$$

where F' is inbreeding of the parent and F inbreeding of the offspring. When F' and F is equal or zero:

$$ab = ½ \quad (5.53)$$

which is the genetic relationship between one of the parents and its offspring.

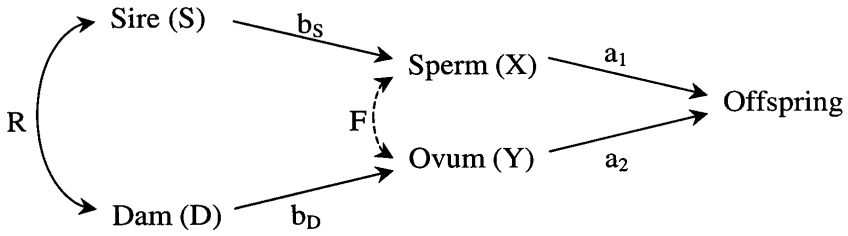


Figure 5.8 Parent offspring relationship. The path from genotype of an individual to its germ cell is b, the path from the germ cell to a fertilized egg is a. The R is the coefficient of relationship between genotypes. The F is the coefficient of inbreeding between the sperm and egg which unite.

The path coefficient is very useful when looking at the relationship between genotype and phenotype and the relationships between traits. The methodology of path coefficients was used by Hazel and Lush (1942) for developing selection index methods in animals. In Figure 5.6, the method of the path coefficients is used to illustrate causes and effects concerning genotype and environment on the phenotype for two traits. (see Lush, 1994, page 710). As Figure 5.5 shows there are directional arrows from G_1 and G_2 to their phenotypes P_1 and P_2, respectively, and likewise from E_1 and E_2 to phenotypes P_1 and P_2, respectively. The arrows have a direction from cause to effect or from G_i and E_i to P_i. The heritability, h^2, is a determination of the phenotypic variation which is caused by the variation in genotype. Thus the path from G_i to P_i is the square root of heritability, which is h_i. Similarly the path from E_i to P_i is the square root of the proportion of the phenotypic variance that is due to the environment.

6. KINSHIP, RELATIONSHIP AND INBREEDING

ANNA K. SONESSON, JOHN A. WOOLLIAMS AND THEO H. E. MEUWISSEN

6.1 Introduction

Inbreeding is a result of mating of related individuals. These related individuals often form a more or less closed population such as a managed breeding population or a wild population that has become isolated from others with little or no migration. In this chapter, we will deal with basic concepts like coefficients of kinship, relationship and inbreeding, which will clarify why relationships are related to the identity of genes, before dealing with the genetic processes that breeding and isolated populations go through.

6.2 Concepts

Malecot (1848) quantified the relationship between two animals U and W by the 'coëfficient de parenté', which is translated in English as the coefficient of coancestry or the coefficient of kinship. The coefficient of kinship is defined as the probability that a randomly sampled allele out of the two homologous alleles of animal U is Identical By Descent (IBD) to a randomly sampled allele of animal W. Two alleles are IBD if they are two copies of the same allele of a common ancestor. If the paternal and maternal alleles of animal U are denoted by u_p and u_m (i.e. those inherited from its sire and dam, respectively) and those of animal W by w_p and w_m, the coefficient of kinship (f_{UW}) is:

$$f_{UW} = ¼ [Prob(u_p=w_p) + Prob(u_p=w_m) + Prob(u_m=w_p) + Prob(u_m=w_m)] \quad (6.1)$$

where $Prob(u_p=w_p)$ is the probability that allele u_p equals w_p.

We can define the kinship of an animal with itself, i.e. f_{UU}, and see that the first and last probabilities in equation (6.1) will be 1.0. In equation form, we can then see that the kinship of animal U with itself is

$$\begin{aligned} f_{UU} &= ¼ [Prob(u_p=u_p) + Prob(u_p=u_m) + Prob(u_m=u_p) + Prob(u_m=u_m)] \\ &= ½ [1+Prob(u_p=u_m)] \end{aligned} \quad (6.2)$$

We can write $f_{UU} = ½ [1+F_U]$, where F_U is the inbreeding coefficient for U, which is defined as the probability that the alleles at a randomly chosen neutral locus are identical by descent (IBD). This is the same probability as the probability that a randomly sampled allele from U's sire is IBD to a randomly sampled allele from U's dam. Therefore, it is clear that the inbreeding coefficient of U is the kinship coefficient between its parents i.e. $F_U = f_{SD,}$ where S and D are the sire and dam of U.

Probability, which always varies between 0 and 1, is used in these definitions, because

T. Gjedrem (ed.), Selection and Breeding Programs in Aquaculture, 73–87.
© 2005 *Springer. Printed in the Netherlands.*

in the absence of complete molecular genetic information we cannot see whether $u_p=w_p$ and have to work with the probabilities. More generally, it is important to recognize that the random sampling of alleles described in the definition of f_{UW} above is done with replacement, which means that even if a parental allele has been sampled once, it is not removed from the pool of parental alleles, but can be sampled again. Random mating, which is one of the assumptions made in deriving many of the standard results in quantitative genetics, also assumes that parents are sampled with replacement. Strictly, random mating should allow selfing, but here we make the assumption that there is no selfing, i.e. the sire and the dam must be different individuals, and that there are two sexes.

Until now, we have looked at IBD at the locus level. However, in real populations, we are generally interested in the average IBD probability over the entire genome. Therefore, all the results will be expectations over all possible loci and populations. From now on, we will focus on this 'quantitative genetic' approach.

An individual can only be inbred ($F>0$) if it traces back through both its parents to a common ancestor so, when there is no selfing, it must have a loop in the pedigree. Inbreeding is inevitable in any closed population over a period of time. To avoid having $F>0$, you need two different parents, four different grand parents, eight different great grand parents, i.e. with the number of distinct ancestors doubled each generation, which is clearly impossible for any finite population. Therefore any finite population will have inbred individuals.

The initial population that is used is often called the base population. Similar to the base population, the initial generation is often called the base generation, whose individuals are assumed unrelated and non-inbred and from which individuals descend. All the alleles carried by individuals in the base generation are considered to be distinct. If two alleles of an individual at a locus have descended from the same allele in the base generation, i.e. they are copies of alleles in the base generation, then these two alleles are IBD as defined earlier. It may be possible to group the alleles in the base generation into forms that are chemically identical in some defined way, e.g. they may share the same sequence of base pairs, or code for the same sequence of amino acids, or some other chemical property. If two alleles in generation t are distinct copies of two base generation alleles, which are identical in form, the alleles are called alike in state (AIS). Being AIS is therefore different from being IBD, since IBD alleles must be AIS (ignoring mutations after the base), but AIS alleles need not be IBD.

For breeding purposes, the base population can be made up of the populations of wild fish that are used to start the breeding program (see Chapter 16 for a discussion on different ways to set up a base population in practice) and the base generation can be the first generation of recorded pedigree. However, in each of the two cases, it is probably not true that all individuals are totally unrelated, i.e. $F = 0$, yet any other coefficient of kinship between individuals is probably also not true. Therefore, the choice of the base generation is quite arbitrary but it is a necessary reference point. It implies that absolute measures of relationship, and thus of inbreeding, within a population only have meaning when the base population is defined, and comparisons between populations are arbitrary

unless the population has a common base population or the information on the population spans a similar time period or generation number. Questions like 'Is a population with an average inbreeding coefficient of 10 % too highly inbred?' cannot be answered without knowing how many generations it took to reach this inbreeding level. However, the rate of increase of the relationship/inbreeding does not depend on the choice of the base population, and it is this parameter that provides a more meaningful measure of what is happening to a population and determines the effective size of the population, which is another important concept that will be introduced in the next paragraphs.

The principal parameter concerned with how fast inbreeding increases is the rate of inbreeding (ΔF). It is defined at time t+1 by:

$$\Delta F = \frac{(F_{t+1} - F_t)}{(1 - F_t)} \tag{6.3}$$

where F_t is the average inbreeding at time t. For populations of constant structure, such as number of parents, means and variances of litter sizes, age structure, selection policy, the expected ΔF in the population is constant over time. The ΔF is a multiplicative, not an additive measure, i.e. the inbreeding coefficient is not expected to increase by a constant amount each generation, but rather a constant proportion of (1-F) is expected to disappear each generation. As will be seen in the next paragraph, (1-F) is directly related to the fraction of heterozygosity that is expected to remain compared to the base generation, so ΔF measures the fraction of heterozygosity that disappears each generation. At the same ΔF, smaller increases in F are expected when F is high than when F is low. The dominator in (6.3) also shows that the maximum inbreeding is 1.0. This is logical because at F=1.0, all individuals have the probability of 1.0 of having exactly the same alleles. When F is very low and ΔF is small the changes in F initially appear to be linear with time i.e. $F \approx t\Delta F$ for small t, and this can be used to estimate ΔF in a population over time. More generally, given a series of average inbreeding coefficients in the population over time, ΔF can be estimated by regression using $\log(1-F_t) = t \log(1-\Delta F) + \text{constant}$.

The term 'effective population size' is often used to describe diversity in populations. The effective population size is defined by $N_e = 1/(2\Delta F)$. Accept no other definition of N_e. N_e individuals of an idealized population with random selection and mating and Poisson litter sizes would have the same rate of inbreeding as the population in question (or approximately ½ N_e males and ½ N_e females randomly selected and mated with Poisson litter sizes). See chapter 3 for a more thorough description of an idealized population.

The widely used additive genetic relationship coefficient for U and W is simply defined as twice the coefficient of kinship: $A_{UW} = 2\ f_{UW}$. This will be called the relationship coefficient in what follows. The relationship of an animal with itself is: $A_{UU} = 1 + F_U$. Because f_{UW} asymptotes to 1.0 as the population becomes more and more related, it follows that A_{UW} asymptotes to 2.0. Therefore relationship coefficients are *not* probabilities. In section 6.3, we will go through how to calculate coefficients of kinship,

relationships and inbreeding.

To sum up what we have looked at so far, Figure 6.1 shows the relationships among the average coefficients of kinship, inbreeding and relationship in a population are shown at a constant rate of inbreeding and random mating when there is no selfing. We can see that (1) all three coefficients start at zero, because individuals are unrelated in the base generation, (2) the coefficient of inbreeding increases with the same rate as the kinship, but starts to increase one generation after the coefficient of kinship, since having two sexes prevents completely random mating, (3) that the coefficient of relationship starts to increase in generation 1, but increases twice as fast as the coefficient of kinship and (4) the coefficients of kinship and inbreeding asymptote to 1.0 and the coefficient of relationship to 2.0.

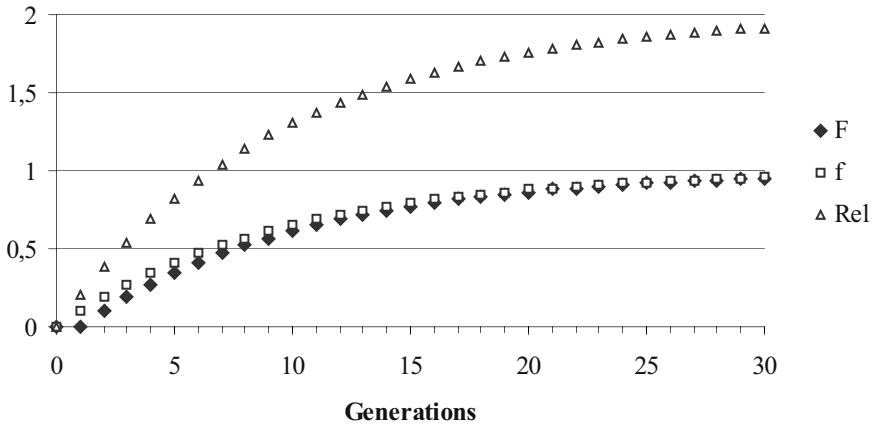


Figure 6.1. Development of coefficients of inbreeding (F), kinship (f) and relationship (Rel) over time.

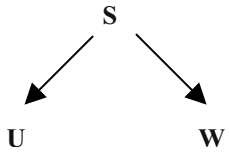
6.3 Calculating coefficients of kinship and relationship

6.3.1 PATHS AND LOOPS

A simple example for the calculation of the coefficient of kinship is that between sire, S, and offspring, U, assuming that the sire and dam of U are both in the base generation, so both $F_U=0$ and $F_S=0$:

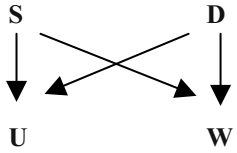
Suppose we randomly choose an allele from U then since F_U =0, it can only be IBD to either of the two alleles carried by S if we sample the allele that had been transmitted by S, and this occurs with probability ½. Since F_S =0, then IBD will occur only if we sample the same allele from S as the one it had transmitted to U, and this probability is also ½, and so $f_{SU} = ½ * ½ = ¼$. It follows that $A_{SU} = 2f_{SU} = ½$.

Next consider two animals U and W that are paternal half sibs with both their dams and their common sire in the base population:



The probability that a randomly sampled allele from U equals its paternal allele, u_p, (i.e., the allele that U inherited from its sire) is ½. Note that we have to sample the paternally inherited allele in order to have any chance of sampling the same allele from W. Similarly, we sample the paternally inherited allele from W with probability ½. Next, we require that S transmitted the same allele to U and to W, which happens with probability ½. It follows that $f_{UW} = ½ * ½ * ½ = 1/8$, and that $A_{UW} = 2f_{UW} = ¼$.

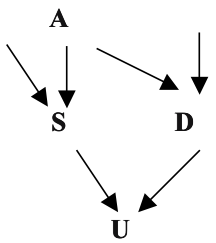
A slightly more complicated example occurs when U and W are full-sibs:



As in the case of the half-sibs, the probability of sampling the paternally inherited alleles of U and W and of those alleles being identical is 1/8. But, here there is also a path from U to W along the dam of U and W. The probability that the maternally inherited alleles of U and W are sampled, and are equal has again a probability of 1/8. Thus the total coefficient of kinship is $f_{UW} = 1/8 + 1/8 = ¼$. And the relationship is $A_{UW} = 2f_{UW} = ½$.

$$g_U = z_{u_p} + z_{u_m}$$

To calculate inbreeding coefficients we can consider all the loops in an individual's pedigree:



If we consider a loop from a common ancestor A, where F_A =0. Then for U to be inbred as a result of this loop, A must pass the same allele to the left-hand side and the right-hand side of this loop, and this happens with probability ½. This same allele must also be passed on along each of the other steps in the pathway S $\rightarrow$ U and D $\rightarrow$ U, with probability ½ for each step. Therefore this loop contributes ½ * ½ * ½ = 1/8; the total inbreeding coefficient would be the sum over all loops using the formula $F_U = \Sigma_{\text{all loops}} (½)^{\text{n\_arrows-1}}$ where n. arrows is the number of arrows in the loop. When A is partially inbred, the probability that the alleles passed to the left and the right of the loop are IBD is ½(1+F_A) and the formula becomes $\Sigma_{\text{all loops}}(½)^{\text{n\_arrows-1}}(1+F_A)$.

The general formulas that emerges from these simple examples is:

$$f_{UW} = ½ \Sigma_{\text{all paths}} (½)^{\text{n\_arrows}} * (1+F_A) \tag{6.4}$$

$$A_{UW} = \Sigma_{\text{all paths}} (½)^{\text{n\_arrows}} * (1+F_A) \tag{6.5}$$

where summation is over all genetic paths that lead from U to W, and n arrows is the number of steps (arrows) in the path from U to W; F_A = the inbreeding coefficient of the oldest ancestor in the path. Note that an individual may be part of 2 or more paths, but the oldest ancestor is different for every path.

6.3.2 THE TABULAR METHOD

There is also a tabular method for calculating coefficients of kinship, which is easier to program into a computer. Suppose animal U (W) inherited allele u_p (w_p) from its sire S and u_m (w_m) from its dam D, then:

$$\begin{aligned} f_{UW} &= ¼ [\text{Prob}(u_p=w_p) + \text{Prob}(u_p=w_m) + \text{Prob}(u_m=w_p) + \text{Prob}(u_m=w_m)] \\ &= ¼ [f_{SW} + f_{SW} + f_{DW} + f_{DW}] \\ &= ½ [f_{SW} + f_{DW}] \end{aligned} \tag{6.6}$$

because Prob(u_p=w_p) is the probability that a randomly chosen allele of S (in this case w_p) equals a randomly chosen allele of B (in this case w_p). Because $A_{UW} = 2f_{UW}$ we have also:

$$A_{UW} = ½ [A_{SW} + A_{DW}]. \tag{6.7}$$

Together with $f_{UU} = ½(1+f_{SD})$ and $f_{UW} = 0$ if U and W are in the base population (or analogously $A_{UU} = 1+ ½A_{SD}$ and $A_{UW} = 0$ if U and W are in the base population), we can use the following algorithm to calculate f (or A) among all individuals, to fill the matrix of coefficients f (or A). We present the algorithm for the **A** matrix, since A is more often used in practice:

1. Sort the animals on age from old to young, so that parents precede their offspring in the list
2. Get A of unrelated base generation animals U and W: A_{UW}=0 and A_{UU}=1.
3. Take the next oldest animal outside the base generation, say U

4. Identify sire S and dam D of U
5. Get A_{XU}, for all animals X older than U, using: $A_{XU}= ½[A_{XS}+A_{XD}]$. This gives the matrix elements of **A**, A(1:u-1; u), and because of symmetry A(u; 1:u-1);
6. Get $A_{UU} = 1+F_U = 1+ ½A_{SD}$. This gives the A(u,u) element of **A**
7. Loop to 3 until all individuals are done

Note that in the above calculation, all the A elements have been calculated before they are needed on the right-hand-side of the equations. If a parent, say a dam, is unknown, it is assumed to be a base generation animal and all the A_{XD} in step 5 and A_{SD} in step 6 are assumed zero.

The extent to which two individuals are related depends on how far we look back in the pedigree, i.e. if we look far enough back all individuals will be related. This is the base generation as discussed earlier (see section 6.2).

6.3.3. PROPERTIES OF THE RELATIONSHIP MATRIX (**A** matrix)

As seen above, the relationships between the individuals of a population can be written in the form of a matrix: the relationship matrix **A**, where each individual has a row and a column in the matrix, and where the (u,w) element is the relationship between individual U and W. For instance, a population consisting of two full sib individuals has the relationship matrix shown below.

$$\mathbf{A} = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix} \tag{6.8}$$

Since the relationship between U with W is the same as W with U, **A** is symmetric. We saw in section 6.2 that the relationship between an individual and itself equals 1, except when the individual is inbred in which case the diagonal of **A**, A(u,u), equals $1+F_u$,where F_i is the inbreeding coefficient of individual i. This implies that the relationship matrix contains the inbreeding coefficients of the individuals. Thus, understanding the relationship matrix and its evolution is one way of understanding inbreeding. Another way is by studying the genetic drift of allele-frequencies at individual loci. Here, we will focus on the **A** matrix method: study the **A** matrix in detail and show its relation to genetic drift.

Let us assume a base population where the effects of the founder alleles are additive, z_i, are all sampled from a distribution with the variance scaled to 1. Any distribution may be used but most realistic is probably the scaled Binomial distribution, e.g. choose randomly one of two alleles. Things are perhaps easier to see when the z_i are sampled from the Normal distribution N(0,1). It may be noted that the additivity assumption can be relaxed by defining z_i as the average effect of the founder allele (averaged over all possible other alleles that can occur together with z_i in an individual). For base population alleles it follows:

$$\text{Variance}(z_i) = 1 \text{ and } \text{Covariance}(z_i;z_j) = 0, \text{ for } i \neq j,$$

and if z_x and z_y are alleles from later generations, which are copies from the same allele in the base population (say z_i) with probability φ_{xy}, then:

$$\text{Variance}(z_i) = 1.0 \text{ and Covariance}(z_x;z_y) = \varphi_{xy}$$

The total genetic value of animal U with alleles z_{u_p} and z_{u_m} is then:

$$g_U = z_{u_p} + z_{u_m} \tag{6.9}$$

with variance :

$$\begin{aligned}&\text{Variance}(g_U)\\ &= \text{Variance}(z_{u_p} + z_{u_m}) = \text{Variance}(z_{u_p}) + \text{Variance}(z_{u_m}) + 2\,\text{Covariance}(z_{u_p}; z_{u_m})\\ &= 2\,(1 + \varphi_{u_p u_m}) = 2\,(1 + F_U) = 4\,f_{UU}\end{aligned}$$

This is because the probability that u_p equals u_m is $\varphi_{u_p u_m}$ but also F_U, and ½(1+F_U) is the kinship coefficient of animal U with itself, f_{UU}. Note that Var(g_U) = 2.0 for non-inbred animals, which is due the assumption that the variance of each allele is 1.0 and thus the variance of the sum of two unrelated alleles is 2.0.

Similarly, the genetic covariance between animals U and W is :

$$\begin{aligned}&\text{Covariance}(g_U;g_W) = \text{Covariance}(g_{u_p} + g_{u_m}\,;\, g_{w_p} + g_{w_m})\\ &+\text{Covariance}(g_{u_p}; g_{w_m}) + \text{Covariance}(g_{u_p}; g_{w_m}) + \text{Covariance}(g_{u_m}; g_{w_p})\\ &+ \text{Covariance}(g_{u_m}; g_{w_m}) = \varphi_{u_p w_p} + \varphi_{u_p w_m} + \varphi_{u_m w_p} + \varphi_{u_m w_m} = 4\,f_{UW}\end{aligned}$$

This is because $\varphi_{u_p w_p}$ equals Prob(u_p=w_p).

The above shows that genetic variances and covariances of the breeding values of animals differ only by a constant factor from their kinship and thus their relationship, where the constant factor, which was 4 here, will depend on the genetic variance of the trait in the base generation.

In general, we may state that the variance covariance matrix of genetic values is:

$$\text{Variance}(\mathbf{g}) = \mathbf{A}\,V_a$$

where **g** = vector of genetic values; V_a = genetic variance.

If the forementioned genetic values are under selection (e.g. phenotypic selection), their variance will be reduced by the selection. Hence, Var(**g**) = **A** V_a holds strictly only if the trait of which **g** is the genetic value (or a correlated trait) is not under selection, but a trait uncorrelated to **g** may be under selection.

The above equality between Var(**g**) and **A** V_a is used in the calculation of BLUP breeding values estimates for **g**. However **g** is often under selection in the case of BLUP-EBV estimation (see chapter 13 for a more thorough description of the calculation of breeding values), but this is not a problem because BLUP uses Var(**g**) = **A** V_a as the a priori distribution of **g,** before the information from the data (and hence the selection) changes the distribution of **g**.

6.4 Consequences of inbreeding

6.4.1 CHANGE OF MEAN AND VARIANCES

Because sampling of alleles is with replacement (see section 6.2) during selection in an idealized population, some alleles get sampled more than once when we start selecting (note the converse, some alleles will also be lost). Some individuals get identical genes in the first generation, and because some alleles are sampled more than once in the first generation some alleles can certainly get sampled more than once in later generations. When we start selecting from the base population, the mean genetic value changes by

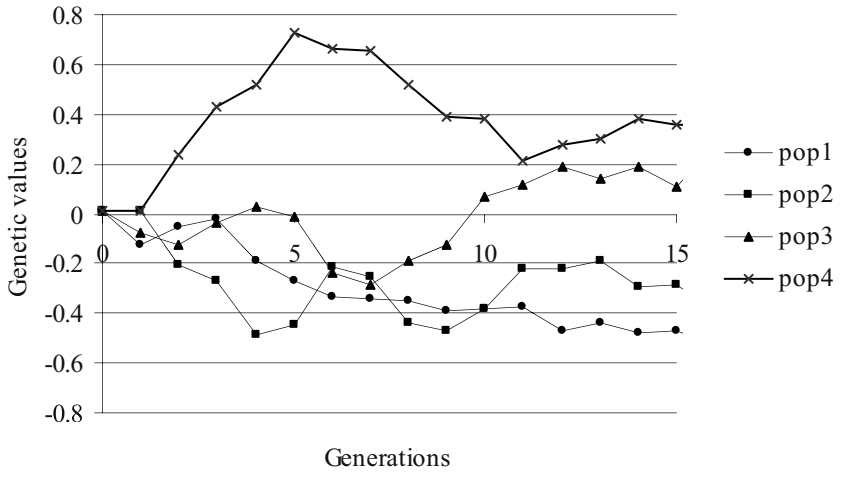


Figure 6.2. Genetic drift over generations.

chance between and within populations through random sampling of alleles alone. In Figure 6.2, we can see that four lines have evolved from one base population. We can

see that the genetic value was zero in the base population, but in two populations the average value has decreased over time and in two it has increased over time. On average over a hypothetically infinite number of populations, the genetic value will remain zero. We can see that the variation between populations has increased. The variation among populations drawn from the same base population is called genetic drift.

In Figure 6.3, we can see that the genetic variation within each line has decreased. Eventually the within line variance will asymptote to 0 as the inbreeding assymptotes to 1. These graphs are made with a very small number of individuals, only ten males and ten females, which make the effects strong. These changes in average values and genetic variation are a direct consequence of inbreeding, which we will see in the following paragraphs.

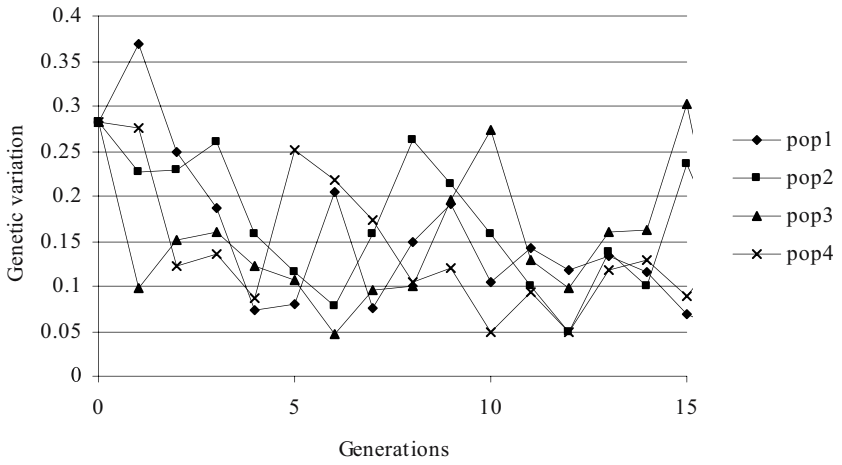


Figure 6.3. Genetic variance within populations over generations.

In matrix form, these effects look like follows:
Let $\mathbf{g}_t$ represent a sub-vector of $\mathbf{g}$ containing the parents of generation t. The genetic mean of the population in generation t is:

$$\text{Average}(\mathbf{g}_t) = \Sigma_i\, g_{ti} /n = \mathbf{1}'\mathbf{g}_t /n$$

where n = number of animals in generation t; $\mathbf{1}$ = n x 1 vector of ones. Now, the genetic drift variance is the variance of the genetic mean:

$$\begin{aligned}
&\text{Variance}(\text{Average}(\mathbf{g}_t)) \\
&= \text{Variance}(\mathbf{1}'\mathbf{g}_t/n) \\
&= \text{Variance}(\mathbf{1}'\mathbf{g}_t)/n^2 \\
&= \mathbf{1}'\,\text{Variance}(\mathbf{g}_t)\,\mathbf{1}/\,n^2 \\
&= \mathbf{1}'\mathbf{A}_{tt}\mathbf{1}\,V_a/n^2 \\
&= 2\,\mathbf{1}'\,\mathbf{f}_{tt}\,\mathbf{1}\,V_a/n^2
\end{aligned}$$

where $\mathbf{f_{tt}}$ = the matrix of kinship coefficients at generation t. Under random mating, including the possibility of selfing, $\mathbf{1'f_{tt}1}$ /n^2 will be the average kinship between the mated animals and thus it will equal the average inbreeding of the offspring in generation t+1, i.e.

$$\text{Variance}(\text{Average}(\mathbf{g_t})) = 2\, F_{t+1}\, V_a$$

This is a well-known relationship between drift and inbreeding. Without selfing and with sexual reproduction, the above relationship is somewhat more complicated, but the rate of increase of drift and inbreeding will show the same relationship.

If the trait is under selection, the genetic variance is reduced. This reduced genetic variance of the selected population can be *approximated* by replacing V_a by $V_{a\infty}$ (1-ρ_∞^2k), which suggests that the drift variance for a trait under selection becomes:

$$\text{Variance}(\text{Average}(\mathbf{g_t}) = 2\, F_{t+1}\, V_{a\infty}\, (1-\rho_\infty^2 k)$$

where ρ_∞ is the correlation between the genetic value and the selection index and k is the variance reduction of the selection index due to selection, which is a function of the selection intensity. Note that the drift variance for a trait under selection is smaller than the drift variance of an unselected trait since $V_{a\infty}$ (1-ρ_∞^2k) < V_a. This is logical because we are managing some of the genetic variance by changing the population mean in the desired direction through selection.

The above shows that the relationship matrix is a very important tool to study genetic drift and inbreeding.

6.4.2 REDUCED HETEROZYGOSITY

The genetic drift leads on a locus level to fixation or loss of alleles, which in turn results in decreased number of heterozygous individuals within a population. When a population is fully inbred (F_t = 1.0), there are no heterozygotes and individuals and populations are either PP (probability p_0) or QQ (probability q_0) i.e. P or Q are fixed within the line. Heterozygotes are expected to decrease within a line in relation to the inbreeding coefficient of the line, such that heterozygosity at time t, $H_t = 2p_0q_0(1-F_t)$. Therefore, the heterozygosity in generation t+1 is related to $(1-F_{t+1})=(1-\Delta F)(1-F_t)$, which is an equivalent representation of ΔF, and if ΔF is constant,
$(1-F_{t_{0+t}})$ = $(1-\sigma\psi\mu\beta o\lambda F)^t(1-F_{t_0})$. This shows more directly how heterozygosity decreases with time: $H_{t_{0+t}} = (1-\sigma\psi\mu\beta o\lambda F)^t H_{t_0}$. For further information, see Crow and Kimura (1970) and Falconer and Mackay (1996). Note that this is an expectation over many lines, but it may not be observed in a single line.

6.4.3 INBREEDING DEPRESSION

Inbreeding depression is only seen in traits with non-additive inheritance, in particular dominance. Inbreeding depression is related to reduced heterozygosity, because for

traits with dominant inheritance we expect *a priori* heterozygotes to be equal or better than one or both of the homozygotes (i.e. d>0). Typical traits with dominant inheritance are fitness traits connected with reproductive capacity or physiological efficiency. These traits can be life-history traits (e.g. fecundity, hatching rate, development rate) and morphological traits (e.g. body size or assymetry) and disease resistance to e.g. recessive diseases where the heterozygotes survive but the homozygotes die (Falconer and Mackay, 1996). Therefore, as the heterozygotes are expected to reduce in frequency due to inbreeding, so we expect to see lower performance in inbred fish, since there are more homozygotes. This is why it is called inbreeding *depression*. In cultured populations with recorded pedigree, inbreeding depression is calculated as a linear regression of the trait value on the inbreeding coefficient of the individuals. The average inbreeding depression of individuals of different breeding populations have been reported at the same average level of inbreeding, e.g. for survival to alevin, inbreeding depression was 11.3 or 4.8 % per ten percent inbreeding in Pacific salmon (Hard and Herschberger, 2002). If the decline is not linearly related to F then this may be indicative of more complex genetic interactions such as epistasis and some studies report such relationships, e.g. survival at an early age in rainbow trout (Kincaid, 1976a,b; Gjerde et al., 1983). Finally, populations of rainbow trout with low rates of inbreeding (Pante et al., 2001a) had lower inbreeding depression in body weight than populations with high rates of inbreeding (Gjerde et al., 1983). This may be explained by the lack of opportunity for new mutations and natural selection to operate.

6.5 Management of genetic variation in fish breeding schemes

In breeding programs the average relationship increases every generation, which cannot be avoided, because no breeding population is of infinite size and therefore it is impossible to find unrelated individuals for more than a certain number of generations. In section 6.4, we saw that inbreeding has negative effects on the genetic make-up of the population. Therefore, inbreeding must be controlled/managed in practical breeding schemes.

- The easiest way to reduce rates of inbreeding is to select more parents per generation. In practical breeding schemes using sib information, this number is limited by the number of full-sib families that the breeding program has. In general, this number should be as large as possible. See Chapter 12 for the effect of different mating ratios between sires and dams, but as a general rule, for a fixed number of parents the more equal the number of sires and number of dams the better. ΔF is most closely related to $1/\min(N_S, N_D)$, where N_S and N_D is the number of selected sires and dams, respectively. A large number of one sex of parent compensates very little for a small number in the other.

- Not only is the number of selected animals important for controlling inbreeding, but also the number of selected animals per family has an effect on the rate of inbreeding. Inbreeding is kept lowest when families contribute equally to the next generation. Especially in populations with large families like fish populations, the risk of selecting many individuals from one family is

large. The most extreme breeding scheme that equalises contributions of families is a within-family selection scheme, where each family contributes a fixed number of parents. Then, animals are selected, because they have the largest deviation from its family mean, implying that they have the best Mendelian sampling term. Alternatively, simply the animals with the highest values are selected within each family. Assuming a constant number of selected parents, walkback selection, as proposed by Doyle and Herbinger (1994) is a kind of within-family selection, where genetic markers are used to determine full- and half-sibs of already selected fish. These sibs will not be selected. The advantage of the walkback selection scheme is that all fish are kept in one large tank, saving costs of keeping many small tanks with full-sib families until tagging. However, as with all within-family selection schemes, walkback selection schemes do not utilise the between family genetic variation component, so that only 50 % of all genetic variation is utilised, resulting in reduced genetic gain.

- When one, or several, unrelated individuals is introduced as a parent in the selection program, inbreeding of their offspring is reduced to zero. Most fish populations have wild ancestors that can be introduced into an on-going breeding program with the purpose to reduce rates of inbreeding of the population. However, with this strategy not only the rate of inbreeding is affected, but probably also genetic advantage over the wild population will be reduced, assuming that the breeding population is better performing for the particular breeding goal than their wild ancestors. In some cases if the selected strain is much better than the wild strain, the offspring from the wild strain may be selected out again in the next round of selection.

- Optimum Contribution is a selection method that maximises genetic response with a constraint on the rate of inbreeding (Meuwissen, 1997; Grundy et al., 1998). The input for this method is a vector of BLUP-EBV, the relationship matrix (**A**) of the selection candidates and the increase in the average relationship that we allow. Note that the restriction is at the level of the kinship or relationship among all of the parents instead of the inbreeding level of the individual offspring. Here is an example to illustrate why the restriction must be at the level of the kinship of the parents: if the restriction would be at the inbreeding level of the offspring, the lowest inbreeding would result when mating an unrelated individual (say a wild migrant) to all candidates. However, in the next generation, all these offspring would be at least half-sibs with an average relationship of 0.25 as we saw in section 6.3, and their offspring will all be grand-offspring from this single migrant. The algorithm calculates the optimum contribution, which can be translated to the number of offspring that each selection candidate should obtain. The method uses a quadratic index to assess the potential contribution accounting for the magnitude of the BLUP - EBV of a candidate and its relationship with the other selection candidates, so that an ideal individual to select is an animal with high breeding value, but that is little related to the other selection candidates. When compared with truncation selection for BLUP - EBVs, optimum contribution selection can

result in up to 60 % higher genetic gain at the same rate of inbreeding. The general explanation for the higher genetic gain is that optimum contribution selection results in increased selection intensity at the same rate of inbreeding. This method is also efficient for multi-spawning populations, where parents and their offspring are selection candidates simultaneously (Meuwissen and Sonesson, 1998; Grundy et al., 2000). The method will either select individuals from the parent generation that are less related, but will also on average have lower BLUP - EBVs, or individuals from the offspring generation, that are more related, but have higher BLUP - EBVs. For these populations with overlapping generations, the ideal individual to select is an old animal with a high breeding value that still has not contributed, i.e. has no offspring. Depending on how stringent the inbreeding constraint is, the increase in accuracy of the BLUP - EBVs as individuals get older, and the number of selection candidates, the method will select individuals either from the parents or offspring generation. The important result is that when compared at the same rate of inbreeding, the use of quadratic indices to optimise contributions provides more gain than truncation selection, mainly as a result of higher selection intensity at the same rate of inbreeding. Therefore, there is no trade-off between gain and inbreeding in their implementation, yet they allow for better management of the genetic resource. (See Woolliams et al., 2002). As ΔF is decreased, providing more secure schemes, so gain may decrease. However, the relationship between ΔG and ΔF is quadratic, so if ΔF is high, then substantial reductions in ΔF result in only small reductions in ΔG.

Without looking on practical limitations on the number of selected parents, non-random mating schemes in combination with optimum contribution selection, e.g. factorial mating (Woolliams, 1989), where all sires are mated to all dams or a mating scheme, where the resulting average kinship of the offspring is minimum, have shown to increase genetic gain with up to 22 % compared with random mating at the same rate of inbreeding (Sonesson and Meuwissen, 2000). With the practical limitations of e.g. a certain number of full-sib families, the superiority of non-random mating schemes will be reduced. In general, any constraint will cost genetic gain.

Optimum contribution selection is a method that is especially useful for breeding nuclea like a normal fish nucleus, because there is a central steering of the use of each selected parent.

We have seen that one effect of inbreeding is the loss of genetic variation and in breeding programs it is this variation that we manage. There are only two processes that increase the genetic variation and make populations survive even with high rates of inbreeding: natural selection and new mutations. An example of natural selection is a selection force that works for an increase in average heterozygosity, such that inbreeding depression is counteracted. Although mutation rate differs over the genome, between individuals, between bases etc, the general variation of mutations in *Drosophila* is about $0.001 * V_e$ for bristle number (Falconer and Mackay, 1996) and $0.005 * V_e$ for body weight in mouse (Keightley, 1998), V_e equals the environmental

variance. No figure has been presented for aquaculture populations. In general, both natural selection and mutation rates are two weak forces in old domesticated populations. However, in recently domesticated species, e.g. many fish populations, natural selection may be important, and should probably be accounted for. More research is needed in this area.

Another relatively open research question is: how high a rate of inbreeding is acceptable in an ongoing breeding scheme? Making some assumptions on the heritability and the coefficient of variation for fitness, and assuming no correlation between production trait and fitness, critical effective population sizes of between 31 and 250 animals per generation were found (Meuwissen and Woolliams, 1994a). As a general rule, effective population sizes of at least 100, i.e. rate of inbreeding of 0.5 %, is often used in practical breeding schemes. It may be noted that inbreeding rate may be critical in aquaculture populations, because they are going from large populations to small during the process of domestication. The time to purge unwanted alleles by genetic drift may then be too short if a too high rate of inbreeding is applied.

7. SELECTION

TRYGVE GJEDREM AND JØRN THODESEN

7.1 Introduction

Selection entails choosing some individuals from the population to produce more offspring than others. These individuals are selected from the animals, which reach sexual maturation, as only those, which can reproduce can affect the future population. Selection does not create new genes, but rather changes gene frequencies. The frequencies of alleles with favourable effects on the phenotype under selection are increased, and the frequency of less favourable genes decreased. In aquaculture, the traits of economic importance in aquaculture are generally quantitative, governed by a large number of genes. The effect of changing the frequencies of the favourable alleles at these genes can be observed as a change of population mean for the trait under selection.

If the purpose of selection is to improve a production trait, the first step is to measure or record this trait on all animals in the population, and then estimate the average and standard deviation. Selection is then practised by selecting those animals, which have highest breeding values.

7.2 Natural selection

The contribution of offspring to the next generation is called the fitness of an individual, adaptive value, or selective value (Falconer and Mackay, 1996) and fitness is the "character" that natural selection selects for. When environmental conditions change, the fitness of existing genotypes may not be optimal any longer. Animals, which have the highest fitness in the current environment, will reproduce at a higher level and have a higher survival rate than less fit animals. The effect of natural selection over generations is to establish a population adapted to the new environmental condition. Thus adaptability is a response of a population rather than of individuals.

If changes in the environmental conditions represent large changes and take place rapidly, a population may be lost or destroyed because natural selection has not been efficient enough to adapt the population to the new conditions. An example showing the inefficiency of natural selection was in recent years demonstrated in South Norway. During a few decades acid rain lowered the pH in freshwater rivers and lakes below the tolerance level of salmonid fish (pH < 5.0). It was shown that tolerance to low pH in freshwater was heritable, $h^2 = 0.09 – 0.33$, (Gjedrem, 1976; Edwards and Gjedrem, 1979) which means that it should have been possible to adapt the fish to a lower pH level. However, acidification occurred too rapidly for the natural selection to adapt the fish to the low pH, and fish populations were lost from hundreds of lakes. It also showed

T. Gjedrem (ed.), Selection and Breeding Programs in Aquaculture, 89–111.
© 2005 *Springer. Printed in the Netherlands.*

that natural selection is rather inefficient over short time periods. One of the reasons for its inefficiency is that it only utilises individual or mass selection.

7.3 Artificial selection

Man creates artificial selection in order to change a population in the wanted direction. However, at the same time the population will continue to be affected by natural selection, which may act in the same or in the opposite direction as artificial selection.

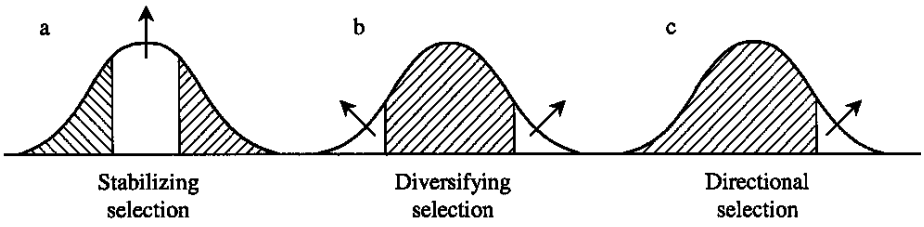


Figure 7.1. Artificial selection may be performed in several ways. a: Animals with average performance are selected. b: Extreme animals are selected and mated inter se. c: Animals with good performance in one direction are selected.

Artificial selection may be performed in several ways and some of them are illustrated in Figure 7.1a. Stabilising selection is practised by selecting phenotypes around the mean and discarding extremes. The aim of stabilising selection is to standardise the population around an average. It results in a fairly constant mean with somewhat reduced variance for the trait in question. For some traits there may be an intermediate optimum which is dependent on the biology and life history of the organism. For example, in fishes an optimum fat content may be wanted in the filet, while high and low fat percentage are less desirable.

In diversifying selection extremes are chosen as breeding animals, Figure 7.1b. If the selected parents were mated at random, no change in average performance is expected and variance in the offspring could be slightly increased. If selected parents in each extreme group were mated, offspring variance would be increased and eventually separate and distinct subpopulations could emerge. This method of selection is rarely used in animal breeding.

The most common form of selection applied in in agriculture and aquaculture breeding programs is directional selection Figure 7.1c. The aim is to improve traits of economic importance. The effect of directional selection for heritable traits is a change in gene frequency at the loci affecting the trait in the next generation. Assuming no change in environmental conditions the average phenotypic value of progenies of selected parents is increased.

Selection against dominant and recessive genes is discussed in Chapter 3 and in section 3.5.

7.4 Prediction of response to selection

The effect of directional artificial selection, which is of particular interest to us, is the change of population mean. This change is called response to selection or genetic gain and is symbolised by ΔG. ΔG is the average superiority of progenies of selected parents compared to the generation before selection. The magnitude or strength of selection applied is called the selection differential S, the distance from the population average to the average of the selected animals, Figure 7.2. Considering that the trait selected for is normally distributed the expected response to selection is:

$$\Delta G = S \cdot h^2 \qquad (7.1)$$

S is expressed in phenotypic units, so must be multiplied by the heritability to express ΔG in genetic units. For example when heritability of the trait selected for is equal to 1 the response to selection will be S and if the heritability is 0 there will be no response to selection.

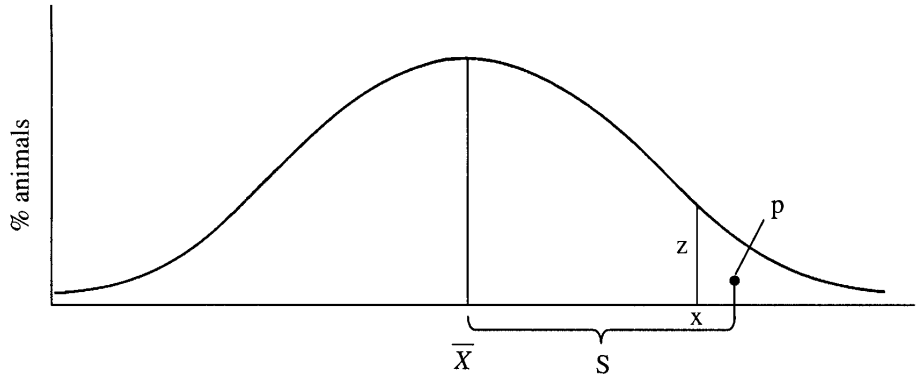


Figure 7.2. Selection differential (S), portion selected (p) and the height of ordinate (Z) at the point of truncation.

If selection differential is expressed in terms of phenotypic standard deviations, σ_P, response to selection may be standardised so selection response can be compared between traits and between populations. The standardised selection differential, S/σ_P, is called intensity of selection, symbolised by i. The standardised selection differential tells us how many standard deviations selected animals are above the population average. The selection differential will be equal to:

$$S = i \cdot \sigma_P \tag{7.2}$$

and the response to selection may we written as:

$$\Delta G = i \cdot h^2 \cdot \sigma_P \tag{7.3}$$

If we want to look response per year, rather than per generation, we must divide by the length of the generation interval in years L:

$$\Delta G/year = i \cdot h^2 \cdot \sigma_P/L \tag{7.4}$$

If we can assume the phenotypes and genotypes of the trait are normally distributed, selection intensity depends only on the proportion of the population selected, Figure 7.2. Let p be the percentage of selected animals which fall beyond the point of truncation (x), and z be the height of the ordinate at the point of truncation, then it follows from the mathematical properties of the normal distribution that:

$$S/\sigma_P = i = z/p \tag{7.5}$$

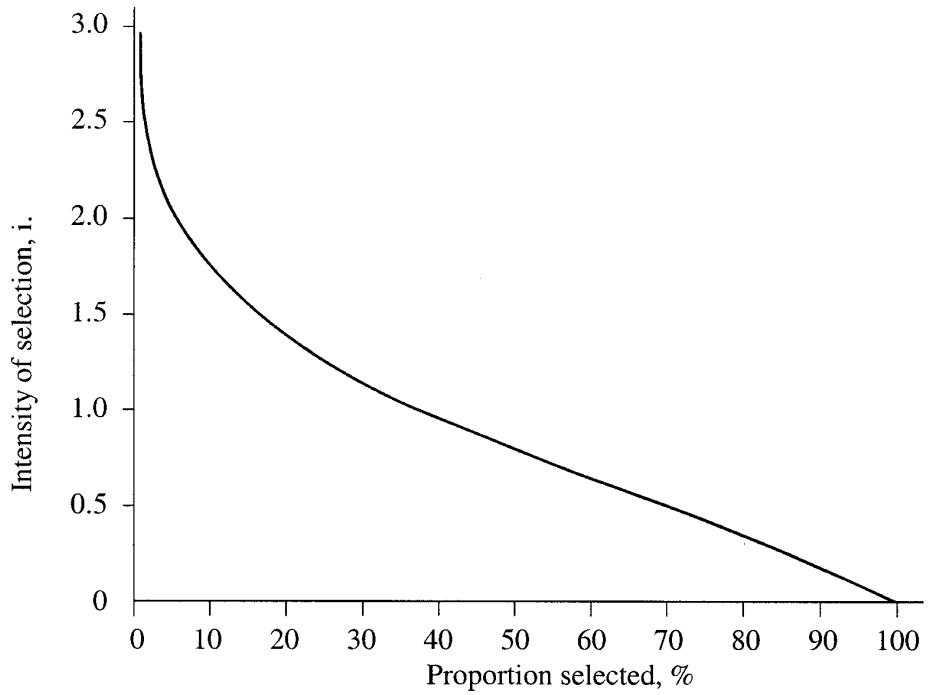


Figure 7.3. Intensity of selection in relation to proportion selected.

The relation between i and percent selected is shown in Figure 7.3. If percent selected animals are known the value of i may be found in Appendix A.

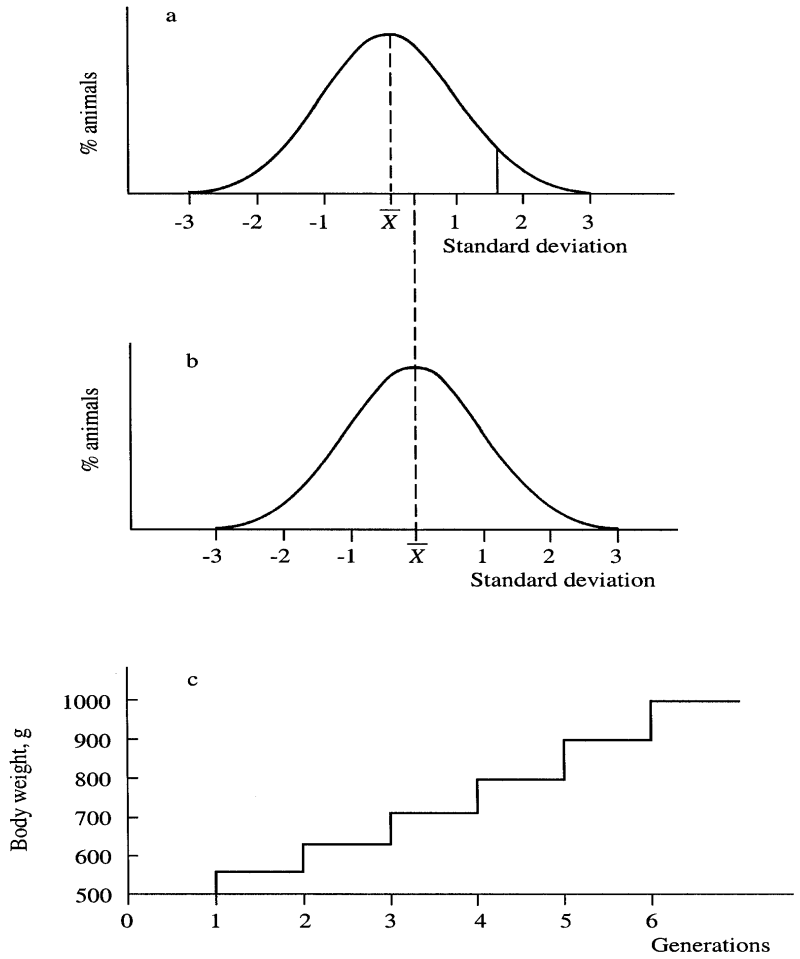


Figure 7.4. Figure 7.4a show the distribution of the parent generation in standard deviation units where 5% were selected as parents. In Figure 7.4b the distribution of the progeny generation is shown together with the change in population average compared with the parent generation. Figure 7.4c illustrates the result of selection for body weight over 6 generations.

7.4.1 PREDICTING GENETIC GAIN, AN EXAMPLE.

Let us consider a population of tilapia with the following parameters:

Body weight $\bar{X}$ = 500 g; standard deviation σ_P = 150 g; heritability h^2 = 0.20; 5 % is selected as broodstock which give a selection intensity of i = 2.063 (from Appendix A).

The base population is illustrated in Figure 7.4a. The response to one generation of selection will be:

$$\Delta G = i \cdot h^2 \cdot \sigma_P = 2.063 \cdot 0.20 \cdot 150 = 61.9$$

Since the standard deviation is given in g the response is 61.9 g. The response to selection in relation to the average is:

$$\text{Genetic gain in \%} = 61.9\text{g}/500\text{g} \cdot 100 = 12.4\ \%$$

The response to selection is cumulative since the average continually increases. If selection is continued for six generations the body weight of the population is expected to double, which is illustrated in Figure 7.4c.

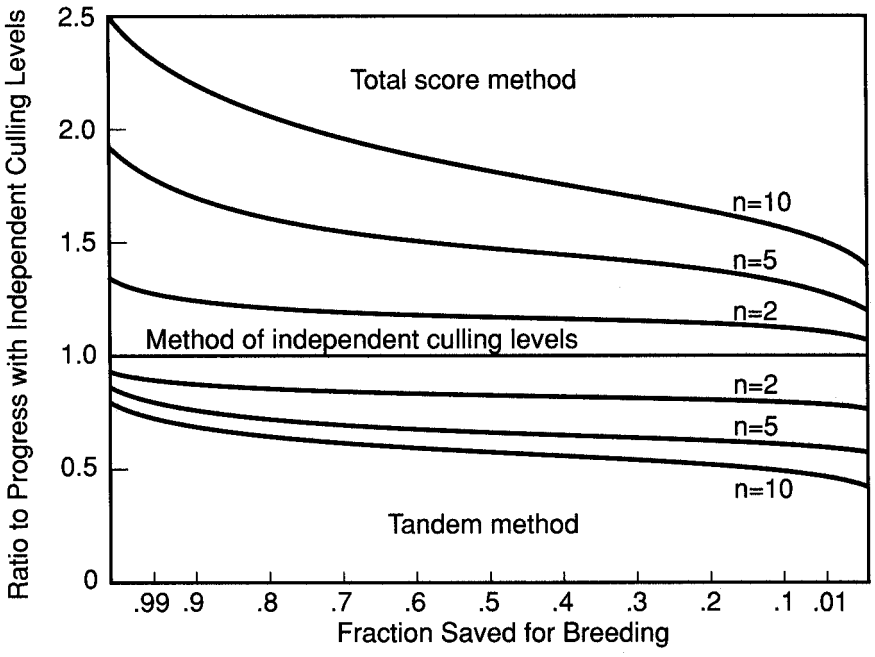


Figure 7.5. Comparison of efficiencies of total-score method and tandem method relative to that of the method of independent culling levels, where the n traits are uncorrelated and are equally important. Reproduced from Hazel and Lush (1942) by permission of Oxford University Press.

7.5 Multiple trait selection

A selection program will usually focus on several traits of economic importance. There are basically three methods of selection when several traits n are involved (Hazel and Lush, 1942).

In tandem selection the individual traits are improved in succession. Selection is carried out first for one trait, until a desired genetic level is reached, then selection for the second, third and following traits are performed.

The second method is to cull simultaneously but independently for each trait, a selection strategy called independent culling levels. Then a level is set for each trait which represent the culling or selection level for each trait.

The third method is to apply selection simultaneously to all traits giving appropriate economic weight, heritability, and phenotypic and genetic correlations between the traits. This method is called selection index or total score.

In Figure 7.5 comparison between the three methods is made. The methods of tandem and total score are related to the efficiency of independent culling level by varying number of traits (n) selected for and percent animal selected for breeding.

The tandem method is the least efficient of the three, while the selection index method is the most efficient in terms of selection response in the direction of the multiple trait breeding goal (Hazel and Lush, 1942). The construction of selection indices for estimation of breeding values is outlined in Chapter 13.

7.5.1 SELECTION FOR A TRAIT AS A RATIO

Selection on a trait defined as a ratio, such as FCR (Feed conversion ratio) or FER (Feed efficiency ratio), involves selection for two traits simultaneously, and as the selection becomes more intense, the selection differential becomes disproportionately larger for

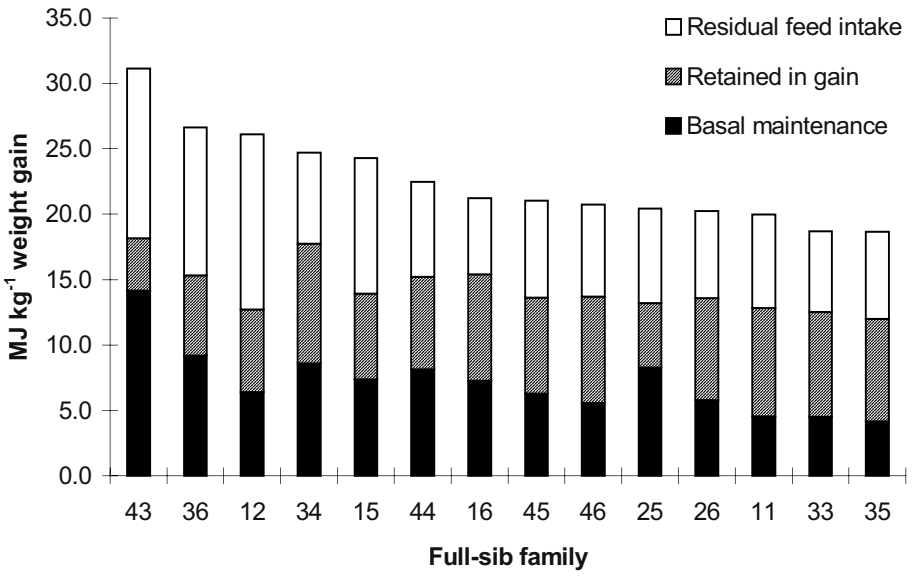


Figure 7.6. Residual feed intake (RFI) calculated for full-sib family groups of Atlantic salmon. Reproduced from Thodesen (1999b) by permission of Department of Animal and Aquaculture Science.

the more variable trait (Gunsett, 1986). Since the coefficient of variation for weight gain generally is approximately twice that of feed consumption (Sutherland, 1965), the selection pressure will be much higher on growth rate than on feed consumption. Alternatively, direct selection can be based on reducing the "Residual feed intake" (RFI) (Figure 7.6), which is the difference between observed feed intake and that predicted on the basis of requirements for production and maintenance of body weight (Kennedy et al., 1993). Such a selection might reduce the energy requirement for body tissue deposition, maintenance and other energy demanding processes (Van Bebber and Meyer, 1994). However, calculation of RFI provides no additional genetic information to that provided by its component traits (Kennedy et al., 1993). Therefore, selection for improved feed utilization should be performed by including the component traits in a selection index with appropriate economic weights.

7.6 Indirect selection

If selection is directed towards one trait (P_1), which is genetically correlated with another trait (P_2), this correlated trait will also be affected by selection. The correlated response (CR) in P_2 depends on the following parameters:

$$CR_{P2} = i \cdot h_{P1} \cdot r_G \cdot h_{P2} \cdot \sigma_{P2} \tag{7.6}$$

where hp_1 and hp_2 is the square root of heritability for trait 1 and 2, respectively, and σ_{P2} is the standard deviation for trait 2. This expression is similar to the expected selection response in P_1 ($\Delta G = i \cdot h^2_{P1} \cdot \sigma_{P1}$) and can be visualised looking at the path diagram in Figure 5.8. When the correlated trait, P_2, is tied to P_1 with a high genetic correlation, it will be changed when selection is directed towards P_1. If the genetic correlation is positive there will be a positive response while the response will be negative when the genetic correlation is negative. The magnitude of the correlated response will also depend on the heritability and standard deviation of P_2.

Indirect selection may be exploited in breeding programs. If a trait is difficult or very expensive to measure or record, a correlated trait may be used instead. The correlated trait may be of importance or may have no economic importance at all.

7.6.1 SELECTION FOR DISEASE RESISTANCE

Fjalestad et al. (1993) discussed the possibility of using correlated or marker traits in order to increase the efficiency of selection for disease resistance in fish. Possible marker traits mentioned are immunological and physiological parameters such as lysozyme (Lund et al., 1995), haemolytic activity (Røed et al., 1990; Røed et al., 1992), cortisol (Refstie, 1982), IGM (Lund et al., 1995; Strømsheim, et al., 1994a), antibody titre (Lund et al., 1995; Strømsheim, et al., 1994b), and plasma α_2 - antiplasmin activity (Salte et al., 1993), which has shown genetic variation and genetic correlations with survival. However, none of these genetic correlations exceeded $\pm$ 0.37, so the correlated response would not be expected to be particularly large.

Growth is also correlated with disease resistance, or at least with survival, Table 5.6. Accordingly, selection for growth should lead to a positive correlated response in survival when selecting for increased growth rate.

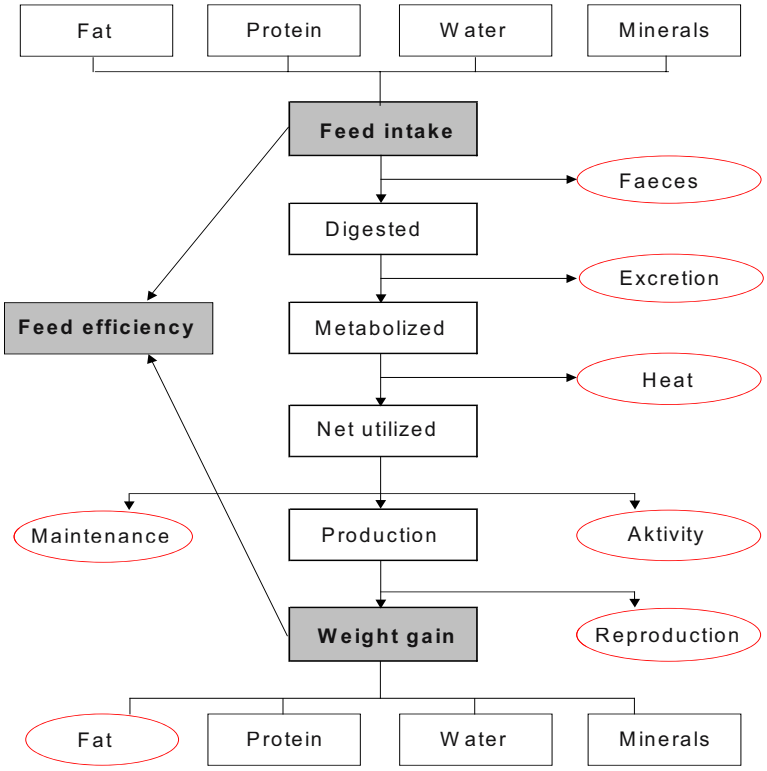


Figure 7.7. Feed efficiency is a complicated trait determined by all traits that influence the relationship between feed intake and weight gain. Selection for a more efficient fish should focus on reducing the expenditure of ingested energy and nutrients.

7.6.2 SELECTION FOR FEED CONVERSION EFFICIENCY

Feed efficiency in fish becomes more important when the production system is intensified, and the feed cost might constitute more than 60% of production cost as observed in intensive Atlantic salmon farming. The feed efficiency in fish farming has been improved both by optimising the feed formulations and improving the feed management. Further improvement of feed efficiency will depend on whether it is possible to improve this trait through selective breeding.

Feed efficiency is a complex trait (Figure 7.7) since it is determined by all traits that influence the feed consumption, loss of nutrients in faeces and excretory products,

expenditure of energy for maintenance and activity, the chemical composition of weight gain or the production of gonads. Feed efficiency in fish is usually expressed as "Feed conversion ratio" (FCR, kg feed consumed per kg weight gain) or, its inverse, "Feed efficiency ratio" (FER, kg weight gain per kg feed consumed). However, these measures are very rough estimates of feed efficiency in fish since they do not allow for changes in the chemical composition of neither the feed nor the weight gain. Alternatively, feed efficiency can be expressed as the retention of protein (g protein in weight gain per g consumed protein) or energy (kJ energy in weight gain per kJ consumed energy), depending on which feed ingredients (sources of protein, fat or carbohydrates) will be restricting in future fish farming.

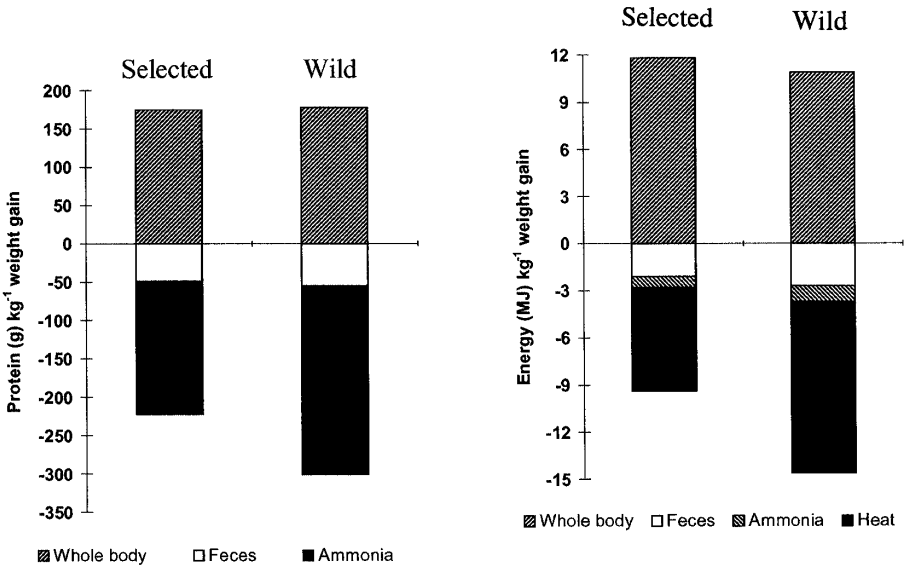


Figure 7.8 Protein and energy budgets of offspring from selected and wild unselected Atlantic salmon. Reproduced from Thodesen (1999b) by permission of Department of Animal and Aquaculture Science.

Selection for increased growth has been used to indirectly improve feed efficiency in fish. Results presented by Thodesen et al. (1999) showed significantly higher relative feed intake, growth and feed efficiency in offspring from Atlantic salmon selected for increased growth for five generations (Selected line) than in offspring of wild unselected salmon (Wild line). Estimated protein and energy budgets (Figure 7.8) indicate that, although fish of the selected line had 17% lower consumption of protein and energy per kilogram gain, they deposited 8% more energy and as much protein as fish of the wild line. Furthermore, fish of the selected line lost a similar amount of undigested protein, but 30% less undigested energy compared with fish of the wild line. The higher relative feed consumption in the selected line, associated with a higher growth capacity, improved feed efficiency by increasing the amount of metabolic energy available for

gain relative to the energy cost for maintenance and activity (McCarthy and Siegel, 1983). This may have contributed to the lower loss of energy as ammonia and heat by the selected line (30 and 39%, respectively) compared with the wild line.

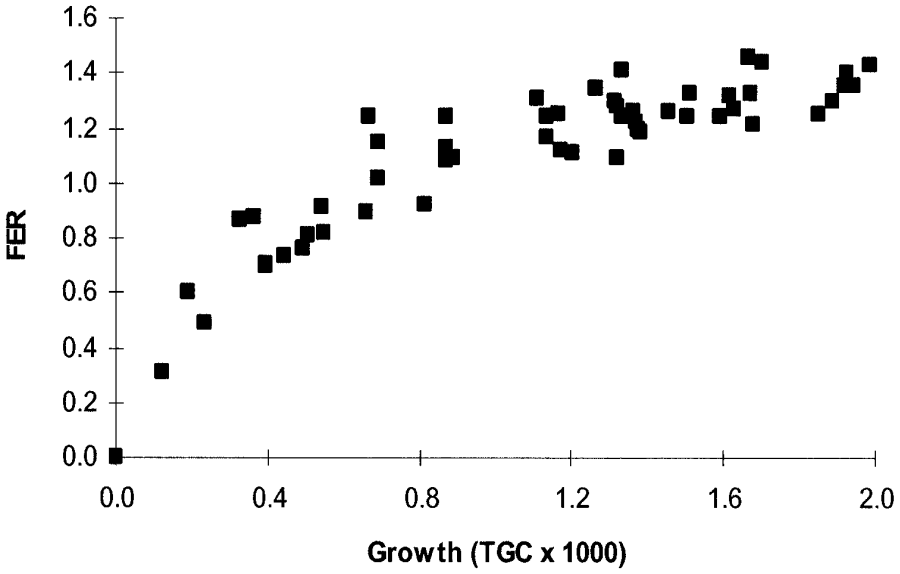


Figure 7.9 Feed efficiency ratio (FER, kg weight gain per kg feed used) in full-sib families of Atlantic salmon with different growth capacity. Reproduced from Thodesen (1999b) by permission of Department of Animal and Aquaculture Science.

Results presented by Thodesen et al. (2001) suggest that the genetic relationship between growth and feed efficiency in Atlantic salmon is non-linear (Figure 7.9), indicating a decreasing response in feed efficiency with increasing growth capacity. Furthermore, Kinghorn (1983b) concluded that young rainbow trout, which consume more feed, grow faster but are not better net converters of food. This implies that at some stage it will become more beneficial to select either directly for improved feed efficiency, as has been concluded for poultry (Sørensen, 1986), or indirectly by selection for reduced fat deposition and/or increased absorption of nutrients.

Silverstein (2003) found significant genetic variation in residual feed intake (RFI) in rainbow trout. There was no correlation between RFI and feed consumption or with growth, but a significant correlation was found with nitrogen retention and with feed efficiency.

In rainbow trout a high correlation, $r_G = 0.78$, is estimated between consumption and growth rate (Gjøen et al., 1993). This estimate is similar to the genetic correlation estimated between growth rate and feed conversion in farm animals, which varies from

0.6 to 0.90 (See Chapter 16). Thus by selecting for high growth rate, feed efficiency will be improved through correlated responses.

Selection for improved feed efficiency may change the requirement of nutrients. The optimal ratio between dietary protein and energy will increase in lean fish compared to fat fish (Åsgård and Hillestad, 1998). Therefore, feed efficiency should preferably be tested using diets formulated for a leaner fish.

7.6.3 SELECTION FOR INCREASED ABSORPTION OF NUTRIENT

Absorption of nutrients can be recorded in individuals or full-sib families of fish reared in the same environment by incorporating an inert marker in the diet (Maynard and Loosli, 1969). Selection for increased absorption of nutrients in fish might improve feed efficiency in a similar way as increasing the nutritional density of the diet, as less feed would be necessary to supply sufficient amounts of nutrients and energy for growth, maintenance and activity. Alternatively, higher absorption may reduce the minimum requirement of essential nutrients (essential amino acids and fatty acids, minerals and vitamins) in the feed.

Results reviewed by Thodesen (1999b) indicate genetic variation in the absorption of energy, protein and amino acids, and minerals in salmonids. The absorption of nutrients may be genetically correlated to fish size. This implies that the absorption of nutrients may be changed in a favourable or unfavourable direction when applying selection for increased growth. The possibility to indirectly improve feed efficiency through selection for increased absorption of nutrients should be followed up with studies where growth, changes in body composition and feed efficiency are also recorded in order to estimate genetic correlations between these traits.

7.6.4 SELECTION FOR REDUCED FAT DEPOSITION

The magnitude of selection response for reduced fat deposition will depend on the amount of genetic variation for this trait. Results presented by Rye and Gjerde (1996) show a genetic variation in deposition of fat in Atlantic salmon. Some of the variation in fat deposition in their study was perhaps due to differences in size and/or growth rate (Rye and Gjerde, 1996; Thodesen et al., 1999a) though Kause et al. (2002) did not find a significant genetic correlation between body weight and fat percentage in rainbow trout. Studies with pigs (reviewed by Vangen and Kolstad, 1986) indicate that leaner individuals may have a higher daily maintenance requirement than fatty individuals. This argues that selection for decreased fat deposition should be combined with selection for increased growth to prevent a reduced feed efficiency due to a higher relative maintenance energy cost. Such a selection will depend on a moderate genetic correlation between growth rate and fat deposition. For Atlantic salmon and other fish species, the magnitude of this correlation is unknown.

7.7 Effect of selection on genetic variance and covariance

The effect of selection is to change gene frequencies, and since genetic variance depends on gene frequencies, it is possible that selection may effect genetic variance. In absence of knowledge of the individual genes affecting quantitative traits, the magnitude of their effects and the frequencies of their alleles, the effect of selection on genetic variance is difficult to predict. One assumption that is often made is that quantitative traits are affected by a very large number of genes, each with a very small effect on the trait. This model is called the infinitesimal model. One prediction that relies on this model is the the Bulmer effect (Bulmer, 1971). The Bulmer effect predicts that in the first generation of selection the genetic variance in the progeny is reduced but the reduction becomes progressively smaller the following generations (Fimland, 1979). It is therefore advised to practise a rather mild selection in the first generation.

According to Falconer and Mackay (1996) the reduction in phenotypic variance (σ^2_P) due to selection is:

$$\sigma^2_{P0}{}^2_P = (1 - k)\,\sigma \tag{7.7}$$

where σ^2_{P0} is the phenotypic variance before selection, k depends on the selection intensity as follows:

$$k = i(i - x) \tag{7.8}$$

where i is selection intensity and x is the distance from population average to the point of truncation. If percent selected animals are known, the value of i and x can be found in Appendix A.

We can re-arrange this formula to predict the effect of selection on the genetic variation:

$$\sigma^2_A = (1 - h^2k)\,\sigma^2_{A0} \tag{7.9}$$

where σ^2_{A0} is the additive variance in the population before selection.

Table 7.2. Response to selection in the second generation relative to that in the first, for some representative values of heritability (h^2) in the base population and of the proportions selected (p %). Reproduced from Falconer and Mackay (1996) by permission of Pearson Education Inc.

	Heritability, h^2				
p %	0.9	0.7	0.5	0.3	0.1
5	0.76	0.79	0.83	0.89	0.96
20	0.78	0.81	0.85	0.90	0.96
50	0.83	0.85	o.88	0.92	0.97

Reduction of genetic variance following selection is particularly high in the first generation of selection and for traits with high heritability. This can be seen from Table 7.2. The reduction of genetic gain in the second generation is because genetic variance is reduced. Heritability will be reduced when the additive genetic variance is reduced and the environmental variation is unchanged.

When selection is directed towards two traits simultaneously, selection may affect the genetic correlation between the traits. Falconer and Mackay (1996) explain the reason for this change: "Those pleiotropic genes that affect both characters in the desired direction will be strongly acted on by selection and brought rapidly toward fixation. They will then contribute little to the variance or to the covariance of the two characters. The pleiotropic genes that affect one character favourably and the other adversely will, however, be much less strongly influenced by selection and will remain for longer at intermediate frequencies. Most of the remaining covariance of the two characters will therefore be due to these genes, and resulting genetic correlation will be negative."

The lesson for us is that we should re-estimate genetic parameters periodically in a selection program to properly account for changes in the genetic parameters, this is particularly important when selection is based on several traits simultaneously.

7.8 Selection response obtained

7.8.1 DISEASE RESISTANCE

One of the earliest documented experiments of selection in fish was that of Embody and Hyford (1925). These authors selected surviving brook trout from a population with endemic furunculosis and increased the survival rate from 2 % in the initial population to 69 % after three generations of selection. Ehlinger (1977) obtained reduced mortality rates due to furunculosis in brown trout and brook trout after selection. Kirpichnikov et al. (1993) reported a selection program against dropsy disease in common carp which started in 1965. Three stocks were involved and the response to mass selection was moderate. Schaperclause (1962) report from a selection experiment in common carp, which showed an average mortality rate of 11.5 % in 65 ponds stocked with progeny of selected fish vs. 57 % in 76 ponds stocked with progeny of non-selected fish. Okamoto et al. (1993) reported on an IPN resistant strain of rainbow trout with average mortality of 4.3 % compared with 96.1 % in a highly sensitive strain.

Fjalestad et al. (1997) obtained a selection response of 12. 4 % for higher survival after challenge test against Taura syndrome in *P. vannamei* although simultaneous selection for growth rate. In Sydney rock oyster Nell and Hand (2003) found a reduction of 22% in mortality after two generation of selection against QX disease. Similar response is reported for selection for resistance to disease in other oysters (Haskin and Ford, 1979; Allen, 1998; Naciri-Graven et al., 1998)

7.8.2 GROWTH RATE

There have been many experiments selecting fish for rapid growth rates. Kincaid et al. (1977) selected for increased body weight at 147 days post-fertilisation. The genetic gain during three generations of selection was 0.98 g or 5 % gain per year. Moav and Wohlfarth (1973; 1976) summaries the results from several mass selection experiments for growth rate in common carp as follow: Selecting for slower growth rate yields response while selection for fast growth rate did not yield positive response. Kinghorn (1983a) reviewed the results from the Israeli selection experiments and concluded that it seems that the report of no response to selection for high growth rate is not conclusive in this case. In a selection experiment to increase growth rate in rohu carp at CIFA, India the response was 30 % in 6 generation of selection (Mahapatra et al., 2004).

In Coho salmon Hershberger et al. (1990) selected for high growth rate in 4 generations and reported a response to selection of 10.1 % per generation. Gjerde and Korsvoll (1999) reported a realised selection differential in Atlantic salmon of 83.9 % during 6 generations or 14 % per generation for growth rate and a reduction of 12.5 units in sexual maturity or 8 % per generation. Gjerde (1986) estimated a response to selection of 13 - 14.4 % per generation in rainbow trout and Atlantic salmon, respectively. Bondari (1983) and Dunham (1987) found a response to selection in growth rate in channel catfish of 20 % and 12 - 18 %, respectively.

In three experiments with tilapia little response to mass selection in growth rate was reported (Teichert-Coddington, 1983; Hulata et al., 1986; Huang and Liao, 1990). However, in a large scale selection project for the Genetic Improvement of Farmed Tilapia (GIFT) in the Philippines to improve growth rate the results were different. The accumulated response the first 5 generations was 85 % with yearly variations between 12 and 17 % (Rye and Eknath, 1999).

Some results from improvement of growth rate from selection experiments and breeding programs with extensive number of animals involved are given in Table 7.3.

Selection experiments have also been implemented in other aquaculture species. Haley et al. (1975) reported that mass selection of adult oysters gave an apparently strong response to selection for growth rate. They concluded that because of large environmental variability, a combination of family and mass selection would be required to achieve maximum response. Newkirk (1980) obtained considerable selection response in growth rate of oysters after one generation of selection. He concluded that 10 -20 % gains per generation in growth rate is a reasonable expectation. Nell et al. (1999) and Toro et al. (1996) report a genetic gain of 9 % and 9-12 % in increased growth rate for oysters, respectively.

Hadley et al. (1991) report a genetic gain of 9 % in growth rate in clam with the first generation of selection. In two selection experiments with scallop Ibarra et al. (1999) estimated the genetic gain to 16 and 18 % gain in total weight.

Table 7.3. Response to selection in growth rate

Specie	x	Percent gain per generation	No Generat-ions	Author
Coho	250 g	10.1	4	Hershberger et. al. (1990)
Rainbow trout	3.3 g	10.0	3	Kincaid et al. (1977)
Rainbow trout	4,0 kg	13.0	2	Gjerde (1986)
Atlantic salmon	4,5 kg	14.4	1	Gjerde (1986)
Atlantic salmon	6.3 kg	14	6	Gjerdeand Korsvoll (1999)
Atlantic salmon	3.5 kg	12.5	1	Flynn et al. (1999)
Channel Catfish	-	12.0-18.0	1	Dunham (1987)
Channel Catfish	67 g	20	1	Rezk et al. (2003)
Channel catfish	491g	17	3	Bondary (1983)
Tilapia	ca. 100g	15	5	Rye and Eknath, 1999
Tilapia		12	12	Bolivar (1999)
Rohu carp	440g	30	6	Mahapatra et al. (2004)
Shrimp	20 g	4.4	1	Fjalestad et al. (1997)
Shrimp	15 g	10.7	1	Hetzel et al. (2000)
Oysters	42 g	9-12	1	Toro et al. (1996)
Oysters	36 g	9	2	Nell et al. (1999)
Oysters		17	1	Newkirk and Haley (1983)
Oysters		20	4	Barber et al. (1998)
Clam	33 g	9	1	Hadley et al. (1991)
Scallop	47 g	16	1	Ibarra et al. (1999)
Scallop		18	1	Ibarra et al. (1999)

In shrimp *(P. vannamei)* Fjalestad et al. (1997) estimated a response in one generation of selection of 4.4 % for growth rate. The low gain obtained can partly be explained by simultaneous selection for resistance against Taura syndrom. Hetzel et al. (2000) found that a direct response to one generation of selection averaged 10.7 % for growth rate in *P. japanicus*.

In general the response to selection for growth rate is very good and higher than obtained for growth rate in farm animals. The main reasons for these differences are:

- In fish and shellfish there is a much higher genetic variance compared to farm animals. According to (Gjedrem, 1998) the coefficient of variation for growth rate is 7-10 % in farm animals while it is 20-35 % in fish and shellfish
- The fecundity is very high in aquatic organisms, which allows much higher selection intensity in aquaculture than in farm animals

7.8.3 FEED CONVERSION EFFICIENCY

There have been no selection experiments selecting for feed conversion efficiency as the primary trait. Correlated responses in FCR following selection for other traits have been observed however. Thodesen (1999a) report a correlated response in feed conversion

when selecting for growth rate in Atlantic salmon. Wild salmon compared with fish from 5th generation selected for growth rate, had 17 % higher intake of energy and protein per kg growth and at the same time 8 % lower retention of both energy and protein. This study demonstrated that fish selected for growth rate have a better utilisation of feed resources compared with unselected animals.

7.8.4 DATE OF SPAWNING

Lewis (1944) reported response to selection in early spawning, egg number and yearling weight. Selection in rainbow trout over 23 years (6-7 generations) is reported to have had a large effect on growth rate, egg production and early spawning (Donaldsen and

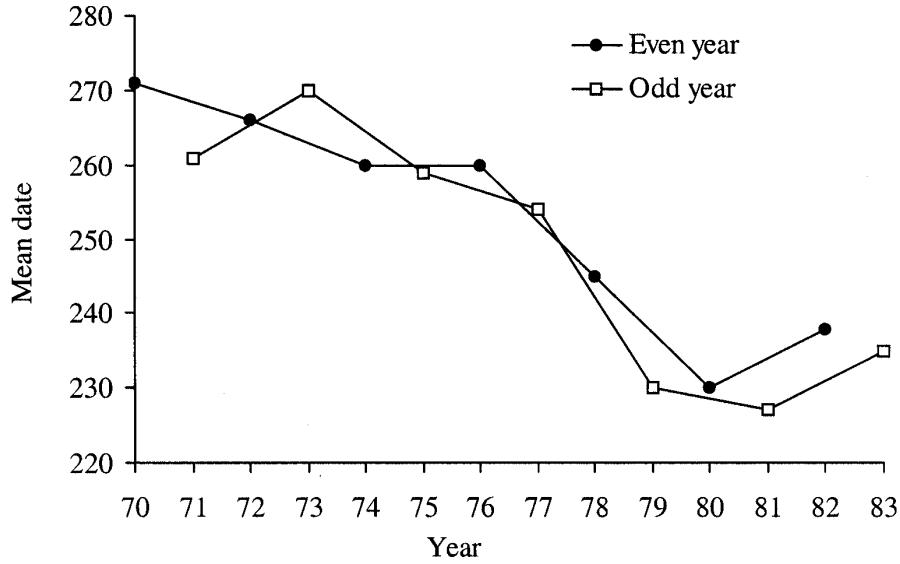


Figure 7.10. Mean spawn date over six generations of selection for early spawn date in two year-classes for the age-2 spawning rainbow trout. Reproduced from Siitonen and Gall (1989) by permission of Elsevier.

Olson, 1955). In both these cases the reported response to selection was confounded with environmental changes because no control was maintained.

Siitonen and Gall (1989) found genetic variation in early spawn date in rainbow trout. They estimated the heritability for early spawn to be 0.55 and 0.53 for two different year-classes. Following selection in the same population, the average response in the selected stock was nearly 7 days per generation through six generations, Figure 7. 10.

7.8.5 SEXUAL MATURATION

In Nile tilapia Longalong et al. (1999) conducted a bi-directional selection experiment for early maturing females. Sixteen pairs of breeders from families with a mean

frequency of 83 % early maturing females (define early maturing) and nine pairs of breeders, all from families with 0 % early maturing females, were selected as parents. More than 3000 fingerlings from the 25 families were tagged and communally stocked in three replicated ponds for testing. The age corrected least square means across ponds and across families within selection groups for frequency of sexually mature females was 57 % and 34 % in progeny of breeders from full-sib families with high and low frequency of early maturing females, respectively. It is proposed that selection for delayed sexual maturation should be included in breeding programs for Nile tilapia.

7.9 Selection limits

Although some long-term selection experiments in laboratory and livestock animals have shown responses continuing over many generations without any sign of approaching a limit, experimental limits to selection response have been demonstrated under certain conditions. Falconer and Mackay (1996) state that selection limits in experimental populations are frequently reached after 20 - 30 generations of selection and that reproduction is often reduced in selection lines. One should distinguish between two situations for selection limit:

1. No genetic variance remains
2. Genetic variance is present but the population fail to respond.

The following are some possible reasons for failure to respond when genetic variance is present:

- Biological or physical limits - egg number per day in laying hens can be regarded as having a biological limit, with the long-lasting eggshell synthesis as the conclusive physical bottle-neck
- Loss of fitness, or antagonistic genetic correlations between selected traits - unfavourable relation between production and reproduction traits in farm animals, or antagonistic relations between production and behaviour traits. An example of the latter is decrease in sexual activity in cocks as a consequence of selection for growth rate in broiler breeds
- Artificial selection may be suspended by natural selection

In several long-term selection experiments the selection response has decreased over generations and finally ceased. Eisen (1980) showed that selection limits in mice were reached after around 30 generations and that response was around 2.4 standard deviations for growth rate and 1.9 standard deviations for litter size. In most selection experiments the response seems to have been higher for productive traits than for reproductive ones, and reproduction appears to have been reduced in lines selected for production traits (Bakken et al., 1998).

Falconer (1960) summarised the results from four selection experiments selecting two ways, two with drosophila (Clayton and Robertsson, 1957; Robertson, 1955) and two with mice (Falconer, 1955; MacArthur, 1949). In these experiments response to

selection had ceased after 30 generations. Number of genes that could account for the observed response varied between 35 and 99.

There are several causes for the reduction in selection response over generations. Selection limits depend very much on number of genes affecting the trait in question. Traits with a few loci will be fixed after a few generations and the result will be no genetic variation and no response to selection. As the number of loci increases it will take more and more generations before genetic variation is effected and limits are reached.

In a closed population there will always be an increase of inbreeding, which will result in fixation of genes and a reduction of the genetic variance of the trait selected for. In addition, as inbreeding increases, fitness in such a population will usually be reduced. Natural selection may act against the artificial selection and thus reduce the genetic gain obtained per generation.

However, if the population under selection is large and the build up of inbreeding is kept low by using many broodstock, at least 50 pairs per generation (Bentsen and Olesen, 2002), selection can continue for many generations without any reduction in genetic gain. This is particularly true when there are a large number of genes contributing to the genetic variation in the trait selected for. Farm animals have been selected for increased production for many generations and there is really no trend of reduced genetic gain.

A good example of a long-term response to selection for a quantitative trait is given by Enfield (1979). He selected for weight of pupa in tribolium for 120 generations without a marked decrease in genetic gain per generation. Both the additive genetic variance and the phenotypic variance have increased during the course of the experiment and the heritabilities have shown only a very slight decline. The total response to selection represents a change of 28 genetic and 17 phenotypic standard deviation units based on the original estimates of these parameters. Figure 7.11a show the development of the control lines and the increase in pupa weight in two selection lines during 120 generations of selection. To illustrate the response to selection in a different manner, the average in the start population and the average in generation 120 together with distribution curves is shown in Figure 7.11b. The distance between the phenotypic averages is 17 standard deviations and there is no overlapping between these populations. In this experiment with tribolium Enfield (1979) showed that it is possible to change a population dramatically by selection, when the breeding population is large, increase of inbreeding is kept low and a considerable number of genes regulate the trait under selection. Estimates of gene number for pupa weight assuming multiplicative effects was found to be at least 170, which the author states as a minimum number.

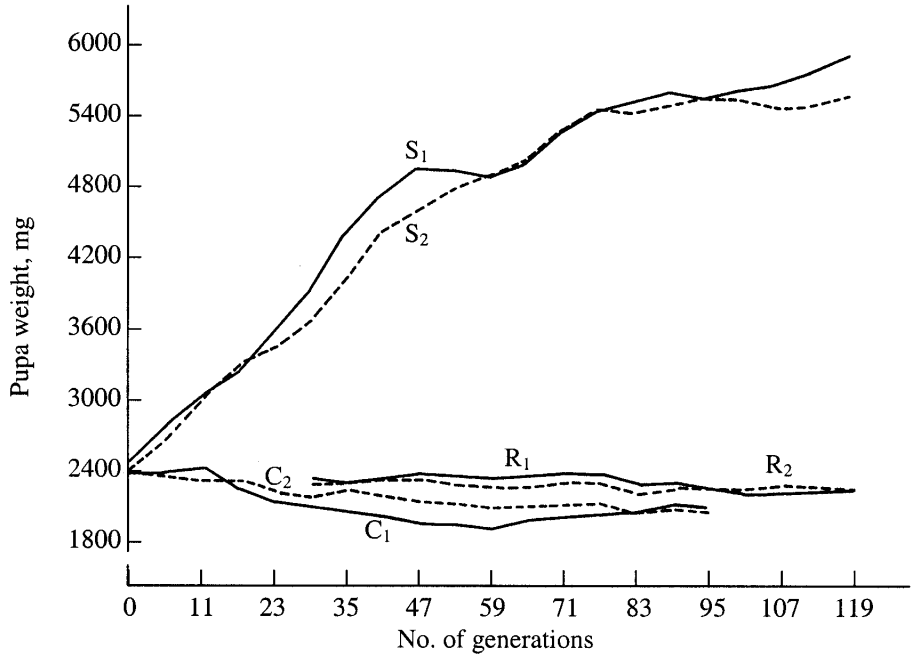


Figure 7.11a. Response to long term selection of weight of pupa in Tribolium, S_1 and S_2 are selected lines for weight of pupa. C_1 ,C_2 ,R_1 and R_2 are control populations. Reproduced from Enfield (1979) by permission of Department of Genetics, Cell Biology and Development.

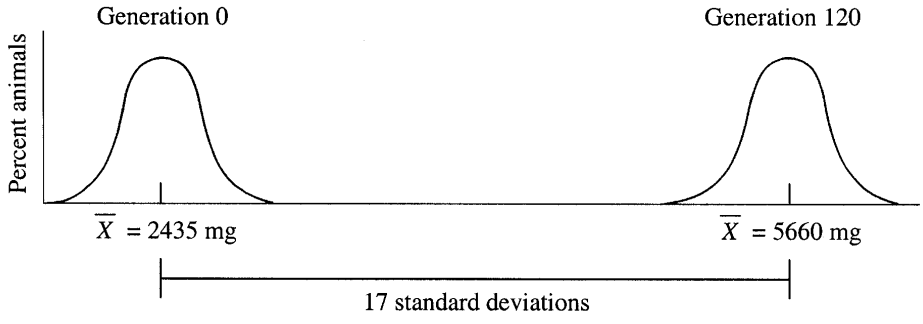


Figure 7.11b. Average and distribution curves for pupa weight for generation 0 and generation 120, the curves are drawn based on data from Figure 7.11a.

7.10 Effect of selection on domestication

Domestication is a process of adapting animals to an environment created by man. Many behavioural, physiological and morphological components of the phenotype can be

changed through selection. In fish and shellfish directional selection is usually applied to improve economically important traits like growth rate, disease resistance, age at maturity and flesh quality. Response to selection is also expected in traits, which are genetically correlated to these production traits. In addition natural selection will adapt the animals to the provided environment. According to Ruzzante (1994) behavioural traits are among the first traits to be affected by the domestication process.

During the early stages of domestication there will be considerable variation in fecundity and variability due to genetic differences in traits which affect the general adaptation of the organism. Such traits are according to Doyle (1983) disease resistance, tolerance to low oxygen and high temperatures, ability to utilize new type of food effectively and so on. Doyle argues further that variation in survival and fecundity is inescapable in the current phase of the domestication of aquaculture stocks and it is obviously important to know whether the variation in commercially valuable traits is positively correlated with fitness. If correlations are positive, then the yield will improve as time passes and the animals become genetically adapted to humane cultivation. If production traits are not positively correlated with fitness, than no change, or a deleterious change, is likely to take place.

It has been argued (Purdon, 1974) that selection for fast growth rate may increase competition among fish. This has been demonstrated when food availability was limited (Swain and Riddell, 1991; Moyle, 1969). However, selection for fast growth rate in medaka showed that when food is available in excess, agonistic behaviour decreased. Under such conditions the dominant and more aggressive fish tended to waste their energy on unnecessary attempts to monopolise the food supply by chasing competitors and thus allow less aggressive fish to grow (Ruzzante and Doyle, 1991). This result is in agreement with Doyle and Talbot (1986a) who concludes that artificial selection for rapid growth will indirectly select for tameness, not aggression. The pattern of decline in aggression with domestication is consistent with that seen for farm animals.

Purdom (1974) showed that the coefficient of variation in growth rate was much higher among fish reared in a group as opposed to individually rearing which indicate that interaction among fish increase variation. It is therefore interesting that Mesa (1991) found that the coefficient of variation for body weight in a wild population was about 4 times larger than in a hatchery population of the same species. This is in agreement with Gjedrem and Fjalestad (1997) who found a high coefficient of variation for body weight of fingerlings of progenies from wild parents of Atlantic salmon, CV = 78, compared with coefficient of variation of body weight for fingerlings of progenies from the 4th generation of selection, CV = 44, Table 7.4. The selection had primarily been based on body weight after 2 years in sea cages. Other marked changes of the fish during 4 generation of selection has according to Gjedrem and Fjalestad (1997) been:

- Distribution of weight of fingerlings is changed from very skewed to become more normally distributed
- The fish is less wild, less aggressive and with a lower level of stress

Table 7.4. Average body weight (x) and coefficient of variation (CV) in four populations of Atlantic salmon parr for four generations of selection for body weight after two years in sea cages. Reproduced from Gjedrem and Fjalestad (1997) by permission of AKVAFORSK.

Generation of selection	Year-class	Populations							
		1		2		3		4	
		x	cv	x	cv	x	cv	x	cv
0	72-75	6.1	78	17.7	75	5.5	84	7.8	75
1	76-79	8.8	67	3.7	55	4.3	74	5.5	59
2	80-83	4.8	59	6.8	48	6.4	58	6.3	64
3	84-87	6.4	50	5.2	40	5.7	51	8.2	56
4	88-91	5.6	43	5.8	47	8.3	43	12.5	42

Domestication is a genetic process, which adapt the animals to the new environment. In an efficient aquaculture production the environmental conditions will continuously be improved which mean that also domestication of the animals must continue. An efficient breeding program will increase the rate of domestication since the animals, which are best adapted will be superior to the others and consequently be selected as broodstock for next generation.

7.11 Socio-economic effects of improved stocks

The adoption of improved Nile tilapia breeds with regard to their impact on equity, the environment and bio-diversity was studied in a project Evaluation of Genetically Improved Tilapia in Asia (DEGITA) (Dey and Eknath, 1996). The production of GIFT and existing Nile tilapia strains were assessed through on-farm experiments in a range of aquaculture systems in Bangladesh, Peoples Republic of China, the Philippines, Thailand and Vietnam. GIFT fish were distributed after two generations of selection.

The results indicated that the cost of production per unit fish produced is lower for the GIFT strain than for the strain of non-GIFT; more than 30 % lower in Bangladesh and the Philippines and about 20 % lower in Peoples Republic of China, Thailand and Vietnam. The estimated yield potential was more than 50 % higher than that of the best existing strains, Table 7.5.

In these countries the majority of the tilapia production is marketed locally and the demand for tilapia is price elastic, both producers and consumers will benefit from technological interventions. As low-priced tilapia is consumed mainly by poor people whose demand is inversely related to price, the major portion of the consumers benefit will go to poor consumers.

Table 7..5 Yield (kg) and production costs (US$) of tilapia farming using GIFT and non-GIFT strains in earthen ponds (per ha.) and cages (per 100 m^2) in five Asian countries. Reproduced from Dey and Eknath (1996) by permission of INFOFISH.

	Bangla-desh	China		Philippines		Thai-land	Viet-nam
	Pond	Cage	Pond	Cage	Pond	Pond	Pond
GIFT:							
Yield	1593	3893	4645	236	1361	2829	743
Cash Costs	463	4191	3548	168	1385	1510	427
Costs of fish/kg	0.29	1.08	0.76	0.71	1.02	0.53	0.58
Non-GIFT:							
Yield	896	3111	4275	153	912	2044	558
Cash Costs	405	4191	3523	168	1375	1517	411
Costs of fish/kg	0.45	1.35	0.82	1.10	151	0.74	0.74
Change in% (GIFT/Non-GIFT):							
Yield	+ 77.5	+ 25.1	+ 8.7	+ 54.2	+ 49.2	+ 38.4	+ 33.2
Cash Costs	14.3	0	0.7	0	0.7	- 0.5	3.9
Costs of fish/kg	- 35.7	- 20.1	- 7.3	- 35.2	- 32.5	- 28.0	- 22.0

8. REPRODUCTIVE TRAITS IN AQUATIC ANIMALS

TERJE REFSTIE AND TRYGVE GJEDREM

8.1 Reproductive cycle

Controlled breeding is only possible if there is enough knowledge about the factors governing reproduction to close the reproductive cycle. Lack of such knowledge has prevented the establishment of breeding programs for several of the farmed species. The elements of controlled breeding, artificial fertilisation and hatchery rearing, must be possible for a breeding program to begin. Artificial reproduction has been achieved for the farmed species listed in Table 8.1.

The reproductive behaviour varies greatly between species. The majority of aquaculture species are season breeders, which are related to climatic conditions like temperature, day length, rainy season and floodwaters. Some species such as salmonids spawn all eggs at once, while other species like flatfishes are batch spawners. The Pacific salmon species spawn once in their lifetime and dye soon afterward while other species spawn once a year and for several years, as do rainbow trout. In the tropics tilapia may spawn repeatedly with 3-4 weeks intervals.

8.2 Sex determination

The basic pattern throughout the animal world is that there are two sexes, male and females. These species are called bisexual. There is, however, in fish considerable variation in sexuality. In some species there are hermaphrodites, in which males and females exist in the same animal. They may first be one sex then the other or both male and female at the same time. Unisexual fish are also known, whereby females produce only female offspring.

8.2.1 UNISEXUAL FISHES

Unisexual fish are rare and none of the fish species farmed is known to belong to this group. Unisexuality occurs either by gynogenese or by hybridogenesis. In gynogenese sperm of a closely related species trigger development of eggs, but do not fuse with it and the result is only female triploid offspring. Natural gynogenese occur in *Poeciliopsis* species and *Carassius auratus gibelio*. Hubbs and Hubbs (1932) were first to described gynogenese in *Poecilia formosa.*

In hybridogenesis, gametic fusion occurs and the paternal genome is expressed, but only the haploid female genome is transmitted to the ovum. Hybridogenese in the genus *Poeciliopsis* has been described by Schults (1971).

T. Gjedrem (ed.), Selection and Breeding Programs in Aquaculture, 113–120.
© 2005 *Springer. Printed in the Netherlands.*

8.2.2 HERMAPHRODITES

According to Purdom (1993) three forms of hermaphrodites are found:

- Protogynous hermaphrodites, in which individuals develop first into females and later into males. This is the most common form, examples are grouper (*Epinepelus tauvina*) and cockoo wrasse *(Labrus bimaculatus)*

- Protandros hermaphrodites, in which the male state differentiates first, examples are sea bream (*Sparus aurata*), oysters (*Grassostrea gigas and Ostrea edulis*). Not all males may undergo sex reversal (Zohar et al., 1978)

- Synchronous hermaphrodites where both male and female states coexist functionally. An example is scallops. This is the least common form of hemaphrodites

Occasionally hermaphrodites may occur in other species, which are normally bisexual. Examples are certain domesticated strains of *Xiphophorus helleri* (Lodi, 1979). Spontaneous hermaphrodites have been reported in brown trout (O'Farrel and Peirce, 1989). Protandros hermaphrodites have also been observed in clams, scallops and abalones.

Sex reversal in hermaphrodite species can occur at almost any time although it mostly occurs after the reproduction season (Brusle, 1987). The mechanism underlying the initiation of sex reversal is still not fully understood, but it seams evident that environmental factors and social behaviour are implicated. Any possible genetic basis for hermaphrodites is not well understood although genetic factors have been suggested.

8.2.3 BISEXUAL FISHES

Most fish species are bisexual with males and females. In most bisexual species some sort of sex chromosome determines sex either through male heterogamety (female XX and male XY) or female heterogamety (female WZ and male ZZ). Sex chromosomes are morphologically indistinguishable from autosomes in most fish species. In some species sex determination has been said to be polygenic and dependent on many genes distributed on many chromosomes, but this last viewpoint has been criticised by Kallman (1965).

8.3 Sex ratio

The relative frequencies of females and males are called the sex ratio. At hatching or early in the life cycle the sex ratio usually is 1:1. Later in the life cycle the ratio may be skewed for several reasons, the survival may be different in the two sexes, in some species age at sexual maturation and following mortality may be different in males and

Table 8.1. Reproduction characteristics.

	Generation interval, years	Hormone for spawning	Egg size, mm	Egg number	Spawning behaviour
Freshwater					
Common carp	2-3	Yes	0.9-1.6	100-150.000/kg	Yearly
Grasscarp	4-7	Yes	1.2-1.4	80-120.000/kg	Yearly
Silver carp	3-5	Yes	1.1-1.4	100-180.000/kg	Yearly
Bighead carp	3-4	Yes		70-120.000/kg	Yearly
Catla	2	Yes		100-200.000/kg	Yearly
Rohu	2	Yes	0.9-1.1	100-300.000/kg	Yearly
Mrigal	2	Yes	0.9-1.1	150-400.000/kg	Yearly
Catfish	2-3	No	1.3-1.6	6-9000/kg	Yearly
Tilapia	0.5-1	No		300-1500	Every 3-4 week
Marine					
Sea bream[1)]	2-3	Yes	0.9-1.1	2-3mill./kg	Several times
Sea bass	2-3	Yes	1.2-1.4	300.000/kg	Yearly/Batches
Grouper[2)]	2-5	Yes	0.7-1.0	300 000/kg	Several times
Yellowtail	3-5	Yes	-	-	
Milkfish	4-5	Yes	0.8-1.2	25-100.000/kg	Several times
Mullet	2-3	Yes	0.6-1.0	1 mill./kg	Yearly/Batches
Turbot	2-3	No	0.9-1.2	1 mill./kg	Yearly/Batches
Atlantic halibut	4-6	No	3	40.000/kg	Yearly/Batches
Wolffish	2	No	5-6		Yearly
Cod	2	No	1.1-1.6	600 000/kg	Yearly/Batches
Crustaceans					
Shrimp (Penae.)	1	Eye ablat	0.2	50-300.000	Several times
Prawn (Macrob.)	1	No	0.6	1 mill	3-4 times/year
Molluscs					
Mussel	1-2	No	0.07	Millions	
Oyster[1)]	1-2	No		1 mill.	
Clams	1	No			
Scallops[3)]	2-3	No	0.01	10-17 mill.	Yearly
Abalone		No		10 mill.	
Anadrome fishes					
Atlantic salmon	3-4	No	6	1600/kg	Once/Twice
Rainbow trout	2-3	No	5	1750/kg	Yearly
Coho	2	No	7	12oo/kg	Once
Chinook	2-3	No	10	800/kg	Once

[1)] Hermaphrodite, Protandrous; Male ⇒ Female
[2)] Hermaphrodite, Protogynous; Female ⇒ Male
[3)] Hermaphrodite, Synchronous; Male and female coexist

females, because of size the difference between sexes grading may also lead to an apparent deviation in sex ratio.

8. 4 Sexual characteristics

Gonads, ovary and testis, together with differences in hormone production are the main primary sexual characteristics in fish and shellfish. In fish the gonads are usually paired organs lying within the coelomic cavity suspended by a dorsal mesentery which supports the ducts for the ovary and testis which open to the outside via a genital pore. With the development of the gonads the animal begins the production of sex hormones which themselves determine the subsequent sexual characteristics and sexual behaviour.

As the animal approaches maturation the gonads grow and fill more and more of the body cavity. Governed by the steroid hormones produced in the gonads, some secondary sexual characteristics develop in some species. Males develop usually more extreme characteristics of shape, colour and aggressive behaviour. Female secondary sexual characteristics are less obvious than those for males.

In general growth rate increases towards sexual maturation. In shoaling or free swimming species like salmonids and mackerels the males typically grow fastest. This is also true for Cichlids. However, in marine flatfishes, several cyprids, eels and sturgeons the females are the largest sex. In shrimp and particular in freshwater prawn the sexual mature male is on average bigger than the female. If sex influences growth rate or size, sex has to be recorded and a correction of body weight for the effect of sex has to be included when the breeding values are estimated.

8.5 Age of maturation

There are big variations between species concerning age of sexual maturation. This has consequences in a breeding program because it sets the length of the generation interval. In Table 8.1 figures are given for average generation interval for some aquaculture species. One should take into account that the generation intervals for the different species depend to a large extent on their environmental conditions and particularly the water temperature. Factors stimulating growth rate will usually shorten generation interval while factors reducing growth rate will lengthen generation interval. For some species sexual maturation occurs before harvest size. In this case later sexual maturation could be part of a breeding goal.

The generation interval or age at sexual maturity varies widely among species. Of farmed species listed in Table 8.1, tilapia has the shortest generation interval and may be sexually mature at an age of 3-5 months of age, but shrimp and prawn also have short generation intervals. Species like grass carp, milkfish, Atlantic salmon and Atlantic halibut have all long generation intervals of 4-5 years. In Atlantic salmon there is a special case where some males become sexually mature as fingerlings (15-20 g).

8.6 Egg and sperm production

8.6.1 PRODUCTION OF EGG

At an early embryonic stage the germ cells are set aside and remain dormant and start development late in organogenesis. When the fish have hatched and start to feed, the gonads develop, the interstitial cells start to produce sex hormones, and consequently the sexual characteristics are developed. Thus the primary oocytes are formed at an early stage of fish life.

The sizes of egg increase from a few micrometers in diameter to several millimetres. The increase in size primarily comprises a huge intake by the oocyte of a glycolipoprotein called vitellogenin, which is synthesised in the liver and controlled by gonadal steroids. After this the eggs are called secondary oocytes and the fish approaches maturation, Fig. 2.7. A parallel process takes place because the release of growth hormone is increased which stimulates the increase of growth rate of the animal.

The number of eggs produced is in general very high and with large variation among species from a couple of hundreds in tilapia to 10 millions in scallops, Table 8.1. This is much higher than for any terrestrial animal species.

The ovulated egg, which has passed the first meiotic division, will have the second meiotic division when the sperm penetrates it and the second polar body is extruded. The egg must be fertilised quickly after addition of water because the eggs swell rapidly in water and this results in the close of the microphyle (a channel leading to nucleus). The time available for carps is about 50-60 seconds.

8.6.2 PRODUCTION OF SPERM

Production of spermatozoa usually takes place during the whole spawning period. For Atlantic salmon and rainbow trout Gjerde (1984b) studied production of semen by stripping the fish at weekly intervals, the Atlantic salmon four times and the rainbow trout three times. The individual variation in the amount of spermatozoa produced was high. The average total volume of semen was 137 ml (20 ml/kg body weight) in Atlantic salmon and 23 ml (5 ml/kg body weight) in rainbow trout. Scott and Baynes (1980) estimated number of spermatozoa to be 15×10^9 per ml of semen in salmonid fishes. The quality of the sperm varies over the spawning period and the viability and ability to fertilise eggs may decline at the end of the spawning period. The mobility of the sperm after they are released into water or ovarian fluid increases for a short time and after a couple of minutes they become immobile and die.

8.7 Fertilisation

Fertilisation takes place quickly after the sperm enter into the water. In order to reach the egg nucleus the sperm passes through the microphyle, a narrow channel through the

outer egg membrane. A relatively large number of spermatozoa is needed to fertilise an egg. Fredrich and Reuter (1982) reported an increase in percentage of fertilised eggs of rainbow trout by increasing the number of spermatozoa from 0.7 to 7.1 millions. For a carp egg 13 000 - 30 000 spermatozoa are required (Pillay, 1990). After a successful fertilisation the microphyle is closed. For anadrome fishes like Atlantic salmon it has been shown that the microphyle is not closed if fertilisation takes place in brackish or full seawater and the egg will die (Billard, 1978).

8.8 Genetic determination of sex

Some species crosses in tilapia produce different sex ratios. Chen (1969) suggested that *O. mosambicus* has a homogametic female (XX) and a heterogametic male (XY). In *O. hornorum* it is reversed with female being heterogametic (WZ) and male homogametic (ZZ). These differences in which sex is the homogametic one can be exploited in tilapia crosses to give unisex populations. Chen (1969) demonstrated that by mating a female *O. mosambicus* with a male *O. hornorum* gave all-male hybrids (XZ). A reversed cross, female *O. mosambicus* and male *O. hornorum* gave three males for every female. Other crosses producing high frequencis of male progenies are:
O. nilotica female and *O. aurea* male; *O. nilotica* female and *O. hornorum* male; *O. mosambica* female and *O. nilotica* male. Lovshin (1980) describe a system to produce all-female hybrids by crossing *O. nilotica* female with *O. hornorum* male. Hulata et al. (1983) proposed a system involving progeny testing to find pairs of broodstock of the species *O. aurea* and *O. nilotica* yielding high frequency of all male progeny.

8.9 Hormone induced sex reversal

It is possible to masculinize or feminize fish. Yamazaki (1983) and Hunter and Donaldson (1983) have reviewed the success rate of various strategies to do this. The most practical approach to obtain the effect is by feeding. By adding sex hormones to the diet from start of first feeding and during some weeks depending on fish species and preparation used, it was possible to get 100 males when a type of testosterone was added to the feed and 100 percentages females when a type of estradiol was added.

Use of hormone for sex reversal may be valuable in some breeding systems for example applying gynogenesis and production of triploids, see Chapter 19. It is also of interest for production of all male and all female populations of fish. Sex reversed genetic females might lack spermatic ducts, and surgery is needed to obtain the sperm.

8.10 Hormone induced ovulation

A common finding in fish and shellfish is that although vitellogenesis is completed in captivity, ovulation does not always occur. This is true for most of the carp species and

many marine fish species. It was therefore a break through for controlled breeding and aquaculture farming in general when hormone treatment was developed. The mammalian luteinizing hormone, LH and human chorionic gonadotropin (HCG) are effective in inducing maturation and ovulation in fishes. Females have to be injected twice and the male once in order to get a good result (Lutz, 2001).

One of the big advantages by applying hormone treatment is that the spawning of many fishes can be synchronised and thus age differences among families will be eliminated.

8.11 Preservation of gametes

Cryopreservation of sperm is widely used in livestock and has been instrumental in the success of breeding programs by progeny testing of bulls in cattle. Investigations in fish by Blaxter (1953) using herring and by Buyukhatipoglu and Holtz (1978) and Stoss (1979) using rainbow trout showed that it was also possible to store sperm in liquid nitrogen and still obtain fertilisation of eggs.

Salte et al. (2004) showed that cryopreservation of sperm can be applied in practice in aquaculture farming. It is a secure way to transfer genes from one country to another without transfer diseases because the milt can be tested for a variety of diseases and parasites beforehand.

Cryopreservation of eggs and zygotes has been tried without much success. The main reason for the difficulty is believed to be the large amount of yolk and cytoplasm in the egg. Small eggs similar to those found in some marine fish and shellfish might be easier to cryopreserve. However, so far there have been no reports of successful cryopreservations of eggs or embryos in any fish species.

Storing of sperm and eggs for several days and in some cases weeks are possible in many aquatic species. For some species protocols for short term storage and transportation of sperm and unfertilised eggs have been developed (Stoss, 1983).

8.12 Influence of reproduction on strategies for genetic improvements

In general aquatic animals have high reproduction potential and a relative low number of breeders are needed for production of larvae and fry. This gives possibilities for high selection intensities and thereby good genetic progress when utilising selection as a tool for genetic improvement. The high reproduction capacity together with very low cost of keeping and replacing breeders also makes it easy to transfer genetic gain from breeding a nucleus via multipliers to the farmers. However the low number of breeders needed in each generation also results in a risk for inbreeding and inbreeding depression in the population. Therefore all selection schemes have to consider the reproductive strategy of the species in question.

There are great differences between aquaculture species in fecundity and reproduction strategies. Important factors such as fecundity under farmed condition and generation interval for the species have to be considered when planning genetic improvement.

Some salmonids spawn once during their life span and in the chase of pink salmon *(Oncorhynchus gorbuscha)* they all spawn when two years old. As a result, there is complete genetic isolation between year classes and development of inbreeding is likely. In this situation cryopreserved sperm could be used to transfer genes between populations, in order to reduce the risk of inbreeding.

For species with long generation intervals and yearly spawning, progeny testing should be considered. In some cases the results from progeny testing are ready when the tested males and females reach sexual maturity, an example is females of Atlantic halibut.

Species that change sex during lifespan gives possibilities to use individuals with high breeding values both as males and females. There is also the possibility to control inbreeding by transferring genes between year classes. In this situation it would be vital to exclude the mating of close relatives, in order to avoid rapid accumulation of inbreeding.

9. METHODS FOR ESTIMATING PHENOTYPIC AND GENETIC PARAMETERS

KARI KOLSTAD

9.1 Estimating genetic variance components from resemblance between relatives

The phenotype (P) is composed of genetic (G) and environmental (E) effects, while the genotype is composed of an additive genetic effect (A), the dominance effects (D) and epistatic effects (I) (Chapter 6) like this:

$$P = A + D + I + E \tag{9.1}$$

The total phenotypic variance is the sum of genetic and environmental variance as stated in chapter 6, while the genetic variance is the sum of the additive variance, and the variance caused by dominance and epistatic effects, i.e:

$$\sigma_P^2 = \sigma_A^2 + \sigma_D^2 + \sigma_I^2 + \sigma_E^2 \tag{9.2}$$

A main objective in breeding and quantitative genetics is to separate the additive genetic variance from the other components in the total phenotypic variance. This can be done by utilising the resemblance between relatives.

9.1.1 PHENOTYPIC COVARIANCES BETWEEN RELATIVES

The knowledge of resemblance between relatives provides a mean for estimating the amount of additive genetic variance in a trait. An understanding of the causes of resemblance between relatives is therefore fundamental to studies of metric characters and to its application in genetic improvement programs in aquatic species and farm animals.

Resemblance between individuals can be expressed using variance components. In the following the composition of covariances between phenotypes of relatives will be shown and used for estimation of additive genetic variance.

Assume that the phenotypic values for two individuals, could be described by the models:

$$P_x = G_x + E_x \tag{9.3}$$

$$P_y = G_y + E_y \tag{9.4}$$

where P_x and P_y are expressed as deviations from the population mean μ. The expected covariances between them can be written as:

$$\mathrm{Cov}[(Gx+Ex)\,(Gy+Ey)]=\mathrm{Cov}(GyGx) + \mathrm{Cov}(EyEx) + \mathrm{Cov}(GxEy) + \mathrm{Cov}(GyEx) \tag{9.5}$$

T. Gjedrem (ed.), Selection and Breeding Programs in Aquaculture, 121–143.
© 2005 *Springer. Printed in the Netherlands.*

which provides:

$$\sigma_{PxPy} = \sigma_{GxGy} + \sigma_{ExEy} + \sigma_{GxEy} + \sigma_{GyEx} \quad (9.6)$$

where σ_{PxPy} is the phenotypic covariance between individual X and Y
σ_{GxGy} is the genetic covariance between the individuals
σ_{ExEy} is the environmental covariance also expressed by σ^2_C
σ_{GxEy} and σ_{GyEx} are the covariance between one of the individual's genotype and the other's environment.

The covariance between the genotypes exists only when the individuals are related to each other. The covariance between the environmental effects for the individuals is different from zero when they are influenced by similar environmental effects and therefore show a common environmental variation. If the common environmental effect is caused by a common dam (full-sibs), then the maternal effect M, will contribute to the environmental covariance by σ_M^2. The covariances $\sigma_{GxEy} + \sigma_{GyEx}$ will exist if the genotype of one of the individuals influence the environment of the other. This may happen in full-sib groups growing up together, and competing for limited feed recourses.

9.1.2 GENETIC COVARIANCE BETWEEN RELATIVES

The proportion of common additive genetic and dominance variation between individuals, i.e. the genetic covariance, is expressed by the additive genetic relationship coefficient, a_{XY}, and the dominance relationship coefficient, d_{XY}, respectively. This connection is essential for describing the genetic covariance σ_{GxGy} between related individuals and for further estimation of the genetic variance. The genotypes of two individuals, X and Y, can be expressed by:

$$G_X = A_X + D_X + (AA)_X + (AD)_X + (AAD)_X + \ldots. \quad (9.7)$$

$$G_Y = A_Y + D_Y + (AA)_Y + (AD)_Y + (AAD)_X + \ldots. \quad (9.8)$$

both expressed as deviations from the population mean, where (AA), (AD) and (AAD) etc. are epistatic effects from interaction between and within additive and dominant effects earlier described by I (Equation 1). Generally the genetic covariance between individual X and Y can then be expressed by:

$$\sigma_{GxGy} = a_{XY}\sigma_A^2 + d_{XY}\sigma_D^2 + a^2_{XY}\sigma_{(AA)}^2 + a_{XY}d_{XY}\sigma^2_{(AD)} + .. \quad (9.9)$$

By knowing the relationship coefficients as shown in Table 9.1, and the variance components for additive, dominance and epistatic effects, the genetic covariance can be estimated. As shown in the following section the contribution of the different genetic variance components to the genetic covariances differ with different types of relationship. Thus, differences in the composition of genetic covariances can be utilised for estimation of genetic variance components.

Table 9.1. Composition of the genetic covariance between relatives including additive, dominance and epistatic effects.

	Genetic relationship coefficients		Proportion of each variance components contributing to genetic covariance							
Relation ship	a_{XY}	d_{XY}	σ^2_A	$\sigma^2_{(AA)}$	$\sigma^2_{(AAA)}$	σ^2_D	σ^2_{DD}	σ^2_{AD}	σ^2_{AAD}	σ^2_{ADD}
Parent-ofspring	1/2	0	1/2	1/4	1/8	0	0	0	0	0
Full-sib	1/2	1/4	1/2	1/4	1/8	1/4	1/16	1/8	1/16	1/32
Half-sib	1/4	0	1/4	1/16	1/32	0	0	0	0	0

9.1.2.1 Genetic covariance between parent and offspring

The additive genetic relationship coefficient, a_{XY}, between parent and offspring is ½ (Table 9.1), while the dominance relationship coefficient is zero. The genetic covariance between parent and offspring will then be:

$$\sigma_{OP} = ½\sigma_A^2 + ¼\sigma_{(AA)}^2 + \quad \sigma_{(AAA)}^2 \ldots\ldots \tag{9.10}$$

i.e. only parts of the additive genetic variance and interactions between additive genetic effects contributes to the genetic covariance between offspring and parent.

9.1.2.2 Genetic covariance between full-sibs

For full-sibs (FS), the a_{XY} is ½ , while the d_{XY} is ¼ (Table 9.1). The genetic covariance will then contain both, additive genetic effects, dominance effects as well as interactions between these:

$$\sigma_{FS} = ½\sigma_A^2 + ¼\sigma_D^2 + ¼\sigma_{(AA)}^2 + 1/16\sigma_{(DD)}^2 + \quad \sigma_{(AAA)}^2 ++ \quad \sigma_{(AAD)}^2 + 1/32\sigma_{(ADD)}^2 \tag{9.11}$$

9.2.2.3 Genetic covariance between half-sibs

For half-sibs (HS), the a_{XY} is ¼ , while the d_{XY} is zero (Table 9.1), i.e. the genetic covariance will contain only additive effects and interactions between them. It is estimated by:

$$\sigma_{HS} = ¼\sigma_A^2 + 1/16\sigma_{(AA)}^2 + 1/64\sigma_{(AAA)}^2 + \ldots \tag{9.12}$$

9.2 Utilising phenotypic covariances between individuals for estimating genetic parameters

In most quantitative traits the genetic variance is mainly caused by additive genetic variance. Consequently, in practical breeding programs it is considered most important to distinguish the additive genetic variance from the total phenotypic variance, and selection programmes are mainly aimed at influencing the additive genetic effects. In

this section, the utilisation of phenotypic covariances between relatives for estimating additive genetic and phenotypic variance will be described.

The complete genotypes for a trait cannot be measured directly. To overcome this problem, observed (statistical) covariances between phenotypes are used to estimate genetic covariances. This is done by estimating variance components on phenotypes from the population under consideration based on a statistical analysis. Statistical variance components estimated from the phenotypes by ANOVA (analysis of variance) based on a statistical model including genetic classes like sire, dam etc. can then be used to derive the biological variance components like the additive genetic variance.

It will be important for the following to understand the difference between the statistical (observed) variance components, which are derived from statistical estimation procedure, and biological variance components, which are related to the effects in the genetic model. The question is under which conditions the covariances between different relatives are valid as estimates of genetic covariances.

In equation (9.6) it was shown that the phenotypic covariance between two individuals is:

$$\sigma_{PxPy} = \sigma_{GxGy} + \sigma_{ExEy} + \sigma_{GxEy} + \sigma_{GyEx}$$

To be able to use the phenotypic covariance as an expression of genetic covariance, it is necessary to assume that σ_{ExEy}, σ_{GxEy} and σ_{GyEx} are not different from zero. If these assumptions do not hold, estimations of genetic covariances will be biased.

9.2.1 ESTIMATING ADDITIVE GENETIC VARIANCE BY UTILISING PHENOTYPIC COVARIANCE BETWEEN PARENT AND OFFSPRING

When phenotypic covariances between relatives are used for estimating genetic covariances, the covariances between environmental effects and between genetic and environmental effects are assumed to be of no significant importance as already pointed out. Then $\sigma_{PoPp} = \sigma_{GoGp.}$ An estimate of σ_A^2 from phenotypic covariance between offspring (o) and parent (p) σ_{PsPo} may then be:

$$\sigma_A^2 = 2\sigma_{PoPp} \tag{9.13}$$

Before it was shown that the genetic covariance, σ_{GoGp} between parent and offspring, is composed by $\frac{1}{2}\sigma_A^2 + \frac{1}{4}\sigma_{(AA)}^2 + \sigma_{(AAA)}^2 + \ldots$Estimating σ_A^2 by multiplying the σ_{PoPp} by 2 will therefore be biased upwards, as it does include parts of the epistatic variance components ($= 2\times(\frac{1}{4}\sigma_{(AA)}^2 + \sigma_{(AAA)}^2 + \ldots)$.

Other sources of bias may also influence the estimates. The offspring may live in the same environment as do the parent causing a $\sigma_{EoEp} \neq 0$, which is often expressed by σ_C^2. This is likely to happen for dam and offspring. Also, an offspring of a sire or dam with genetic high ranking may receive extra care causing a $\sigma_{GoEp} \neq 0$. Phenotypic covariance between dam and offspring will often include a common heritable maternal effect, as

the maternal effect the mother provides to her offspring is half of the maternal effect the grandmother provided to the mother (σ_M). This results in the phenotypic covariance between dam and offspring:

$$\sigma_{PdPo} = \sigma_{GdGo} + \tfrac{1}{2}\sigma_M^2 + \sigma_C^2 \quad (9.14)$$

Estimating σ_A^2 as from the phenotypic covariance between offspring and dam by $2\sigma_{PdPo}$ will be generally more biased than estimating σ_A^2 from phenotypic covariance between offspring and sire by 2σPsPo. In many fish species, the offspring receive no more care from their mother after hatching, and the maternal effect may be of less importance compared to farmed animals. Gall (1974) found that growth of rainbow trout fingerlings was affected by age of dam and egg size. Chevassus (1976) found maternal effect in early growth of rainbow trout. However, it disappeared after two months.

If the parents are selected, it is not recommended to use the covariance between parent and offspring for estimating genetic variance, as selection will reduce variation between the parents' phenotypes. This will automatically lead to a reduced covariance between parent and offspring. Estimating the genetic variance as twice the genetic covariance will then result in an underestimation of the true genetic variance.

9.2.2 ESTIMATING ADDITIVE GENETIC VARIANCE BY REGRESSION OF PHENOTYPIC VALUE OF OFFSPRING ON PHENOTYPIC VALUE OF PARENTS.

The relationship between parent and offspring can be utilised for estimating genetic variation by using the regression of offspring on the parent's phenotype avoiding bias caused by selection.

$$b_{O|P} = \frac{\Sigma(x_i-x)(y_i-y)}{\Sigma(x_i-x)^2} = \frac{\sigma_{PoPp}}{\sigma_P^2} \quad (9.15)$$

where x_i and y_i is the phenotypic value of the parent and offspring respectively, and x and y are the means for parent and offspring population respectively.
The statistical model will be: $Trait_O = b_{OP} * Trait_P + e$. A regression analysis based on this model will provide an estimate of the regression coefficient b_{OP}. The expected value of this regression when the parent is the sire, and we assume $\sigma_C^2=0$ and $\sigma_{GxEy}=0$, is:

$$E(b_{O|S}) = \frac{\sigma_{GsGo}}{\sigma_P^2} = \frac{\tfrac{1}{2}\sigma_A^2 + \tfrac{1}{4}\sigma_{AA}^2 + \ldots}{\sigma_P^2} \quad (9.16)$$

The expected value of this regression when the parent is the dam, where $\sigma_M^2 \neq 0$ and $\sigma_C^2 \neq 0$, is:

$$E(b_{O|D}) = \frac{\sigma_{GdGo}}{\sigma_P^2} = \frac{\tfrac{1}{2}\sigma_A^2 + \tfrac{1}{4}\sigma_{AA}^2 + .. + \tfrac{1}{2}\sigma_M^2 + \sigma_C^2}{\sigma_P^2} \quad (9.17)$$

The σ_P^2 is the phenotypic variance for the population, which can easily be estimated. Then σ_A^2 can be estimated by:

$$\sigma_A^2 = \sigma_P^2 \times 2b_{O|S} \quad (9.18)$$

or

$$\sigma_A^2 = \sigma_P^2 \times 2b_{O|D} \quad (9.19)$$

These estimates of genetic variance are not influenced by selection among parents. The assumptions for using this method are: 1) offspring must remain unselected, and 2) no assortative mating.

The bias by using phenotypic covariance between offspring and parent for estimating genetic variance related to epistasis, environmental covariance and covariances between genetic and environmental effects is however still present.

The regression of offspring on sire is preferred for estimating genetic variation in some species like Tilapia, as σ_M^2 is not present, and as it might be easier to avoid σ_C^2 and σ_{GxEy}. The estimation will increase in accuracy with increasing number of paired observations.

9.2.3 ESTIMATING ADDITIVE GENETIC VARIANCE BY UTILISING PHENOTYPIC COVARIANCE BETWEEN SIBS

Covariance between sibs provides more flexible ways of estimating genetic variance compared to covariance between parent and offspring. Some advantages are that it is present in the same generation, and that it also includes individuals without offspring, which is important for traits like slaughter traits.

9.2.3.1 Half-sib analysis

In fish breeding as well as in breeding of farmed animals it is common that sires are mated to more than one dam. If each dam receives only one offspring, the offspring's phenotype can be described by the statistical model:

$$Y_{ij} = \mu + s_i + e_{ij} \quad (9.20)$$

where s_i is the random effect of sire and e_{ij} is the random effect of the offspring ij. The expected value of s_i and e_{ij} is $E(s_i)=E(e_{ij})=0$. The expected values of s_i^2 and e_{ij}^2 are $E(s_i^2)=\sigma_s^2$ and $E(e_{ij}^2)=\sigma_e^2$. The phenotypic covariance between two half-sibs is then:

$$Cov((Y_{il} - \mu)(Y_{i2} - \mu) = Cov((s_i + e_{i1})(s_i + e_{i2}) = \sigma_{Si}^2 + \sigma_{Siei2} + \sigma_{Siei1} + \sigma_{ei1ei2} \quad (9.21)$$

The phenotypic covariance between individuals based on *genetic terms* from section 2 is:

$$\sigma_{P1P2} = \sigma_{G1G2} + \sigma_{G1E2} + \sigma_{G2E1} + \sigma_{E1E2} = ¼\ \sigma_A^2 + \quad \sigma_{AA}^2 + \ldots \quad (9.22)$$

Assuming that the half-sibs are kept in different environments and not environmentally influenced by each other's genetic value, the last three covariances are set to zero. The genetic covariance between half-sibs is, as derived earlier, equal to:

$$\sigma_{G1G2} = \sigma_{Si}^2 = ¼\,\sigma_A^2 + \quad \sigma_{AA}^2 + \ldots \qquad (9.23)$$

i.e. from a genetic point of view this is a covariance between half-sibs, while from a statistical point of view, this is the variance between sires. This also provides an estimate of within half-sib group variation, which is the difference between total phenotypic variation and the covariance between half-sibs.

The analysis of variance for a balanced design based on the model $Y_{ij} = \mu + s_i + e_{ij}$ is shown in Table 9.2.

Table 9.2. Formula for calculating sums of square, mean squares and expected mean squares for a half-sib analysis for balanced data.

Source of variation	DF	Sum of squares	Mean SS	Expected mean sum of squares
Between sires	s-1	Sire SS	MSs=Sire SS/(s-1)	$\sigma_e^2+n\sigma_s^2$
Within sires	s(n-1)	Error SS	MSe=ErrorSS/(s(n-1))	σ_e^2
Total	sn-1	Total SS	Total SS/(sn-1)	

When the mean squares are equated to their expectations, the estimated variance components are as follows:

$$\text{Between-sire variance } \sigma_s^2 = (\text{MSs-MSe})/n \qquad (9.24)$$

$$\text{Residual variance } \sigma_e^2 = \text{MSe} \qquad (9.25)$$

The variance σ_{Si}^2 is equal to the variance caused by the sires in the population if the sires in the sample are a random sample of sires in the population, and if the dams are picked out at random and randomly mated to the sires. If the sampled sires are selected, the half-sib analysis will provide an underestimation of σ_A^2. If the estimate of σ_s^2 is biased, the estimate of σ_A^2 will be four times this bias. The remaining covariances can be set to zero if no special treatment is offered to offspring from particular sires (σ_{Siei1}=0), and offspring with a common sire do not have a common environment (σ_{ei1ei2} =0).

9.2.3.1.1Example

From a total of 3240 Tilapia, harvest weight was recorded. Variance components were estimated from ANOVA based on a sire model. The ANOVA follows, together with estimates of expected mean squares.

General Linear Models Procedure
Dependent Variable: WEIGHT

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	56	3830495.08	68401.70	34.39	0.0001
Error	3183	6330271.09	1988.78		
Corrected Total	3239	10160766.17			

R-Square	C.V.	Root MSE	WEIGHT Mean
0.38	23.93	44.60	186.35

Source	DF	Type III SS	Mean Square	F Value	Pr > F
SIRE	53	623837.03	11770.51	5.92	0.0001
POND	1	1285.37	1285.37	0.65	0.42
SEX	1	1980876.05	1980876.05	996.03	0.0001
AGE	1	684745.54	684745.54	344.31	0.0001

Source	Type III Expected Mean Square
Sire	Var(Error) + 59.34 Var(SIRE)

Then σ_s^2 = (MSsire – MSerror)/59.34 = 164.84
and $\sigma_A^2 = 4\,\sigma_S^2 = 659.37$

9.2.3.2 Full-sib and half-sib analysis in a hierarchical design
In breeding programs for aquatic species, as for several farm animals, each dam has several offspring, and the offspring's phenotypes is better described by the model:

$$Y_{ij} = \mu + s_i + d_{ij} + e_{ijk} \tag{9.26}$$

where each sire is mated to more than one dam, and s_i, d_{ij}, and e_{ijk} are the effect of sire, dam within sire and the offspring within sire and dam, respectively. The expected values of s_i, d_{ij} and e_{ijk} are $E(s_i)=E(d_{ij})=E(e_{ijk})=0$.The expected values of s_i^2 ,d_{ij}^2 and e_{ijk}^2 are $E(s_i^2)=\sigma_s^2$, $E(d_{ij}^2)=\sigma_d^2$ and $E(e_{ijk}^2)=\sigma_e^2$.

Phenotypic covariances between full-sibs may be utilised for estimating genetic variation. With a proper mating design this covariance can be used for estimating dominance effects and maternal effects as well as additive genetic effects.
The covariance between to full-sibs is:

$$\begin{aligned} \mathrm{Cov}((Y_{ij1} - \mu)(Y_{ij2} - \mu)) &= \mathrm{Cov}((\mu + s_i + d_{ij} + e_{ij1})(\mu + s_i + d_{ij} + e_{ij2})) \\ &= \sigma_s^2 + \sigma_d^2 + 2\sigma_{sd} + 2\sigma_{se} + 2\sigma_{de} + \sigma_{e1e2} \end{aligned} \tag{9.27}$$

σ_{sd} is set to zero if there have been no assortative mating and the parents are not related. σ_{se} and σ_{de} are set to zero if offspring have not been treated according to parents genetic merit. σ_{e1e2} is set to zero if there are no covariance between environmental effects except those caused by the fact that they belong to the same litter, which is already included in the σ_d^2.

When the mean squares are equated to their expectations, the estimated variance components are as follows:

Between-sire variance σ_s^2 = (MSs-MSd)/nd
Between-dam variance σ_d^2 = (MSd-MSe)/n
Residual variance σ_e^2 = Mse

$$\sigma_s^2 = ¼\ \sigma_A^2 + \sigma_{AA}^2 + \ldots \tag{9.28}$$

$$\sigma_d^2 = ¼\ \sigma_A^2 + ¼\ \sigma_D^2 + \sigma_{AA}^2 + \ldots + \sigma_M^2 + \sigma_C^2 \tag{9.29}$$

Table 9.3. Formula for calculating sums of square, mean squares and expected mean squares for a combined full-sib and half-sib analysis in a hierarchic design.

Source of variation	DF	Sum of squares	Mean SS	Expected mean sum of squares
Between sires	s-1	Sire SS	MSs=Sire SS/(s-1)	$\sigma_e^2+n\sigma_d^2+nd\sigma_s^2$
Within-siresand between dams	s(d-1)	Dam SS	MSd=Dam SS/(s(d-1))	$\sigma_e^2+n\sigma_d^2$
Within dams	sd(n-1)	Error SS	MSe=Error SS/(sd(n-1))	σ_e^2
Total	sdn-1	Total SS	Total SS/(sdn-1)	

The covariance between full-sibs is then:

$$\sigma_s^2 + \sigma_d^2 = \tfrac{1}{2}\sigma_A^2 + \tfrac{1}{4}\sigma_D^2 + \tfrac{1}{4}\sigma_{AA}^2 + \ldots + \sigma_M^2 + \sigma_C^2 \qquad (9.30)$$

The variance components are estimated according to Table 9.3.
The additive genetic variance can be estimated in three ways from σ_s^2 and σ_d^2.

1) $\sigma_A^2 = 4*\sigma_s^2$ which is equal to $\sigma_A^2 = 1/a_{HS} * \sigma_s^2$
2) $\sigma_A^2 = 4*\sigma_d^2$ which is equal to $\sigma_A^2 = 1/(a_{FS}-a_{HS}) * \sigma_d^2$
3) $\sigma_A^2 = 2*(\sigma_s^2+\sigma_d^2)$ which is equal to $\sigma_A^2 = 1/a_{FS} * (\sigma_s^2+\sigma_d^2)$

The three alternatives are connected with different degree of bias as described earlier. Using maternal half-sibs for estimating additive genetic variance (alternative 2)) provides the most biased estimate as it includes the dominance effects, maternal effects and common permanent environmental effects within full-sib family. The importance of these three effects may however vary between traits and species of fish. Example 9.2.2 from lice attack in Atlantic salmon show, that the contribution of these effects results in minor differences in the estimates of additive genetic variance when using dam or sire component for estimation. A lower estimate when using the dam component, indicate no important maternal or dominant component.

9.2.3.2.1 Example

From a total of 2049 Atlantic salmon, sea lice attack was recorded. Variance components are estimated from ANOVA based on a nested model, where dam is nested within sire according to the hierarchic mating design usually used in salmon. The ANOVA table follows together with estimates of expected mean squares and additive genetic variance components.

The GLM Procedure:
Dependent Variable: Total lice

Source	DF	Sum of Squares	Mean Square	F -Value	Pr > F
Model	296	1217.19	4.11	1.57	<.0001
Error	1749	4575.28	2.62		
Corrected Total	2045	5792.47			

R-Square	Coeff Var	Root MSE	Total lice Mean
0.21	80.22	1.62	2.02

Source	DF	Type III SS	Mean Square	F Value	Pr > F
Sire	159	595.41	3.7447382	1.43	0.0006
Dam(Sire)	136	405.88	2.9844053	1.14	0.1352
Weight	1	89.69	89.6917045	34.29	<.0001

Source	Type III Expected Mean Square
Sire	Var(Error) + 6.6 Var(Dam(Sire)) + 12.1 Var(Sire)
Dam(Sire)	Var(Error) + 6.5 Var(Dam(Sire))
Weight	Var(Error) + Q(Weight)

σ_e^2 = MSE = 2.62
σ_d^2 = (2.98 – 2.62) / 6.5 = 0.055
σ_s^2 = (3.74 – 6.6*0.06 – 2.62) / 12.1 = 0.062

Estimates of additive genetic variance are then:
$\sigma_A^2 = 2\ (\sigma_d^2 + \sigma_s^2) = 0.234$
$\sigma_A^2 = 4\ \sigma_d^2 = 0.22$
$\sigma_A^2 = 4\ \sigma_s^2 = 0.248$

9.3 Mixed model types of analysis

Estimation of variance components with balanced data is done by equating the mean squares to the corresponding expected mean squares before solving the equations by simple sequential substitution as shown before. Experimental data is often unbalanced, and is influenced by fixed as well as random effects. The fixed effects need to be corrected for to provide unbiased estimates of genetic and phenotypic parameters. The theoretical framework for the analysis of unbalanced data under mixed models equation (MME) has been described by Henderson (1953). The basic model is:

$$\mathbf{y = Xb + Za + e} \tag{9.31}$$

where

y = n x 1 vector of observations; n = number of records
b = p x 1vector of unknown fixed effects; p= number of levels for fixed effects
a = q x 1vector of unknown random animal effects; q = number of levels for random effects
e = n x 1 vector of unknown random residual effects corresponding to the elements of **y**
X = a known design matrix of order n x p, which relates the elements of **y** to those of **b**
Z = a known design matrix of order n x q, which relates the elements of **y** to those of **a**

Both **b** and **a** may be partitioned into one more factors depending on the situation.

It is assumed that the expectations **(*E*)** of the variables are ***E*(Y) = Xb**;
***E*(a) = 0**; ***E*(a) = 0**; ***E*(e) = 0** and that the residual effects, which include random environmental and non-additive genetic effects, are independently distributed with variance σ_e^2. Therefore, var(**e**) = $\mathbf{I}\sigma_e^2 = \mathbf{R}$; var(**a**) = $\mathbf{A}\sigma_a^2 = \mathbf{G}$ and cov(**a,e**)=cov(**e,a**)=0, where **A** is the additive genetic relationship matrix. Then

$$\begin{aligned} \text{var}(\mathbf{y}) = \mathbf{V} &= \text{var}(\mathbf{Za} + \mathbf{e}) \\ &= \mathbf{Z}\,\text{var}(\mathbf{a})\mathbf{Z'} + \text{var}(\mathbf{e}) + \text{cov}(\mathbf{Za},\mathbf{e}) + \text{cov}(\mathbf{e},\mathbf{Za}) \\ &= \mathbf{ZGZ'} + \mathbf{R} + \mathbf{Z}\text{cov}(\mathbf{a},\mathbf{e}) + \text{cov}(\mathbf{e},\mathbf{a})\mathbf{Z'} \\ &= \mathbf{ZGZ} + \mathbf{R} \end{aligned} \quad (9.32)$$

This statistical model is the basis for both estimation of breeding values as seen in Chapter 3, and for estimating genetic and phenotypic parameters. Some advantages are that all pedigree information can be utilised for estimating genetic parameters, and that solutions for fixed and random effects for an unbalanced design is provided simultaneously.

To achieve unbiased estimates, inclusion of a complete relationship matrix is required. With unbalanced data, due to non-orthogonality between the factors, there are a variety of methods each with a unique set of statistical properties. Which method to choose will depend on the model, data layout, properties desired and ease of computation. After the introduction of MME by Henderson, the methods for estimating variance components have evolved through efforts by many scientists resulting in several methods. Frequently used methods for estimation variance components are REML (Restricted Maximum Likelihood), MIVQUE (Minimum Variance Quadratic Unbiased Estimators) and MINQUE (Minimum Norm Quadratic Unbiased Estimators), which are all based on maximum likelihood estimations.

The only common property shared by most used methods is that they do not depend on the magnitude of fixed effects in the same model. Some of the methods are unbiased, while others are nearly unbiased. Methods differ in the magnitude of sampling variance of estimates. Some methods require prior knowledge or guesses about the parameters that one is trying to estimate, and other methods do not. Some methods are iterative in nature and require more computing time than methods that are not iterative.

To estimate variance components from MME a model must be specified including specifications of fixed and random effects. From the specified model, the data and the pedigree information, matrixes are designed to provide the estimates. The computer software will from the specified model and the chosen estimation method solve for the estimates of the unknown components after defining quadratic forms, and then equate the quadratic forms to their expected values under the rules of the method, where the expected values are functions of the unknown components of variance defined by the method chosen. Finally, it will solve for the estimates of the unknown components. The last step is done by inverting a matrix and multiplying the inverse with the computed quadratics, or by iteration involving repeated calculations.

9.4 Estimation of heritability

9.4.1 ESTIMATING HERITABILITY FROM REGRESSION AND CORRELATIONS BETWEEN RELATIVES

When genetic and phenotypic variance is estimated for a population, the heritability can easily be estimated. As described before, resemblance between relatives is utilised for

the estimation of genetic variation in a trait. The genetic and phenotypic variances are then used for estimating heritabilities. Most frequently used is the regression of offspring on parents and sib correlations (Table 9.4).

Table 9.4. Methods for estimation of heritabilities

Regression offspring-parent:	$b_{OP} =$	$\frac{Cov(OP)}{\sigma_P^2}$	$=$	$\frac{\frac{1}{2}\sigma_A^2}{\sigma_P^2}$	$= \frac{1}{2} h^2$
Regression offspring-mid-parent:	$b_{OP} =$	$\frac{Cov(OP)}{\frac{1}{2}\sigma_P^2}$	$=$	$\frac{\frac{1}{2}\sigma_A^2}{\frac{1}{2}\sigma_P^2}$	$= h^2$
Half-sib correlation:	$r_{HS} =$	$\frac{Cov(HS)}{\sigma_P^2}$	$=$	$\frac{\frac{1}{4}\sigma_A^2}{\sigma_P^2}$	$= \frac{1}{4} h^2$

For the regression of offspring on parents, the data is obtained in the form of measurements of parents (one or the mean of both), and the mean of their offspring. The cross-product of the paired values provides an estimate of the covariance. If both parents are measured and included in the regression, the regression will be an estimate of the heritability. If the variance differs between the two sexes, the mid-parent value cannot be used, and the heritability must be calculated separately for each sex. Estimates based on offspring-parent relationship will be influenced by common environmental effects if they are kept in the same unit. Further, the maternal effect will if present, influence the estimates. Both these situations result in estimates biased upwards.

Estimating heritability based on sib-correlations assumes an unselected population for reasons discussed before. Further, it demands a mating structure, where each male or female is randomly chosen and randomly mated to more than one randomly chosen female or male. The individuals measured thus form a population of full-sib and half-sib families.

Generally, the half-sib correlation and the regression of offspring on father are the most reliable ways to estimate heritabilities. As pointed out earlier, full-sib correlation is least reliable of all as they are influenced by common environmental effects, and dominance variance.

The presence of a common environmental effect will vary between species. As seen in example 9.2.3.2.1 Atlantic salmon in that experiment did not show important effects of common environment for the trait lice attack. Tilapia, where the dam keeps the eggs and newly hatched fry in her mouth, may be more influenced by a common environmental effect.

9.4.1.1 Example

Data from harvest weight in tilapia from a NORAD (Norwegian Agency for Development Cooperation) project at RIA 1 (Research Institute in Aquaculture No.1) in 2002 illustrates this.

HWT $\sigma_{d^2(Y)}=169.54$ $\sigma_{A^2(Y)}= 4\sigma_{s^2(Y)} = 424.84$

$\sigma_{s^2(Y)}=106.21$ $\sigma_{E^2(Y)}= \sigma_{e^2(Y)} - 2\sigma_{s^2(Y)} = 1536.9$

$\sigma_{e^2(Y)}=1875.98$ $\sigma_{P^2(Y)}= \sigma_{d^2(Y)} + \sigma_{s^2(Y)} + \sigma_{e^2(Y)} = 2151.73$

Heritability estimated from the sire-component is h^2= 4* (106.21/2151.72)= 0.197. Estimating heritability from the dam-component would over-estimate the heritability: h^2=4* (169.45/2151.73) =0.315. This implies a maternal effect, providing a common environmental effect to full-sibs, influencing the data.

9.4.2 COMBINING INFORMATION BY REML

With pedigree records available, it is possible to estimate the heritability from all sorts of relatives. Estimation methods based on a restricted maximum likelihood (REML) routine combining estimates from different relationships, suitably weighted, to provide the best estimate of heritability. Estimations in REML are done by iteration based on a mixed model. The iterative process involves two steps. 1) the value of λ, a derivation of the heritability, is calculated from estimates of genetic and phenotypic variances, and the mixed model equations are solved for the fixed effects and the breeding values. 2) updated estimates of the genetic and phenotypic variances are determined from the predicted breeding values. Steps 1) and 2) are repeated until the estimates converge.

REML accounts for selection, and all available pedigree information is utilised.

9.4.3 HERITABILITY OF FAMILY MEANS AND WITHIN-FAMILY DEVIATIONS

Family design in breeding programs connected to aquaculture are characterised by large family groups. The selection in properly designed breeding programs using individual tagging is mainly based on family selection, within family selection or a combination of these two. It is therefore of significant relevance to focus on the between family and the within family component of variation and how heritabilities of family means and within family deviations can be estimated.

We are now concerned with the variation of the observed family mean, $\sigma^2_{B,}$ and the observed within-family deviations, σ^2_W. Under between family selection the source of information is the mean phenotypic value for the members of a family, P_f, and the phenotypic variation of the family mean is expressed as:

$$\sigma^2_{Pf} = \frac{1 + (n-1)t}{n} V_P \qquad (9.33)$$

where t is the intra-class correlation between the full-sibs within the family expressed as σ^2_B/σ^2_T, and σ^2_T is the total variance which is equal to V_P.

The additive genetic variance of the family mean is then:

$$\sigma^2_{Af} = \frac{1 + (n-1)r}{n} V_A \qquad (9.34)$$

where r is the correlation between breeding values within family (eg. ½ for full sibs, ¼ for half-sibs). The observational components of between family variation are however:

$$\sigma^2_f = \sigma^2_B + (1/n)\, \sigma^2_W \qquad (9.35)$$

Under within family selection the source of information is deviation between members within each family. Phenotypic variation for the within family deviation is expressed as:

$$\sigma^2_{Pw} = \frac{(n-1)(1-t)}{n} V_P \tag{9.36}$$

where t is the intra-class correlation between the full-sibs within the family. The additive genetic variance of the within family deviation is then:

$$\sigma^2_{Aw} = \frac{(n-1)(1-r)}{n} V_A \tag{9.37}$$

The observational components of within family variation are however:

$$\sigma^2_{w} = \sigma^2_{W} + (1/n)\, \sigma^2_{W} \tag{9.38}$$

Estimating h^2 of family means and of within family deviation is then:

$$h^2_f = \frac{1 + (n-1)r}{1+(n-1)t} h^2 \tag{9.39}$$

and

$$h^2 w = \frac{1 - r}{1 - t} h^2 \tag{9.40}$$

9.5 Estimation of repeatability

When more than one measurement of the trait can be taken on each individual, the phenotypic variance can be partitioned into variance between and within individuals. The between-individual variance is composed of the genetic variance, σ_G^2, and a special environmental variance, σ_{Es}^2, while the within-individual variance is expressed by the general environmental variance, σ_{Eg}^2. The repeatability is then expressed by:

$$r = \frac{\sigma_G^2 + \sigma_{Es}^2}{\sigma_P^2} \tag{9.41}$$

while

$$\frac{\sigma_{Eg}^2}{\sigma_P^2} = 1 - r \tag{9.42}$$

The repeatability expresses the proportion of the variance of a single measurement that is due to genetic and specific environmental effects, effects which are then expected to be repeated in equal strength in repeated measurements. Assumptions when estimating repeatability are 1) equal variances in the different measurements, and 2) that different

measurements reflect what is genetically the same trait. The repeatability will indicate gain in accuracy from multiple measurements. More will be gained by repeated measurements with a low repeatability. The phenotypic variance of a single measurement is:

$$\sigma_P^2 = \sigma_G^2 + \sigma_{Es}^2 + \sigma_{Eg}^2 \tag{9.43}$$

The phenotypic variance from the mean *n* repeated measurements within an individual is then:

$$\sigma_{P(n)}^2 = \underbrace{\sigma_G^2 + \sigma_{Es}^2}\ + \underbrace{1/n\ \sigma_{Eg}^2}$$

$$= \quad r\,\sigma_P^2 \quad + \quad \frac{(1-r)}{n}\sigma_P^2 = \left[r + \frac{(1-r)}{n}\right]\sigma_P^2 \tag{9.44}$$

Increasing the number of measurements reduces the contribution from the general environmental variance to the total variance for each measurement. The effect is however less if the repeatability is high. The advantage of taking repeated measures can be expressed by the ratio of $\sigma_{p(n)}^2$ on σ_p^2, and how this ratio decreases with increasing numbers of measurements when the repeatability is less then 1.

In a breeding programme, the advantage of repeated measurements is seen from the gain in accuracy in breeding evaluation. The phenotypic variance is reduced when σ_{Eg}^2 is minimised, and the additive genetic variance increases in proportion to phenotypic variance.

Phenotypic variance and heritabilities may be estimated on basis of means taken from family groups, which is how traits are measured in family groups of fish before tagging. This can be handled in the same way as repeated measurements within individuals. The variance is then partitioned into within and between family groups. The additive relationship between individuals within a family group determines which value *r* will be given. For a half-sib group, r = ¼ , with phenotypic variance for the mean estimated as:

$$\sigma_{P(n)}^2 = \left[\frac{1}{4} + \frac{(1-\frac{1}{4})}{n}\right]\sigma_P^2 \tag{9.45}$$

9.6 Phenotypic and genetic correlations between traits

9.6.1 ESTIMATING CORRELATIONS FROM COVARIANCES BETWEEN RELATIVES

Correlations between traits are a measurement of the degree of common factors causing the variation in each of two traits. The genetic correlation in particular is of great

importance in a breeding program. The genetic correlation provides an understanding of the genetic background of each trait, and is important to provide more accurate predictions of breeding values. Further, it is important to know if and how much the improvement in one trait will cause simultaneous changes in other traits. With correlations close to 1 or –1 one can expect that selection for only one of the traits gives a high correlated response for the other trait. If two traits are unfavourably correlated, an improvement of one trait will lead to an unfavourable change in the other. It may then be challenging to achieve genetic improvement in both such traits within a population, or even to change one trait without changing the correlated trait. Knowledge about genetic correlations between production traits like growth and fitness traits are important to prevent or minimize reduced fitness as a consequence of efficient modern breeding programmes.

As seen in Chapter 5, correlation between traits is composed by the covariance between the traits and the standard deviation for each of the traits. As for the variance of individual traits, the phenotypic correlation has both genetic and environmental causes, and the phenotypic correlation will fully reflect the genetic correlation only when the environmental correlation is zero.

$$r_p = h_X h_Y r_G + (1-h^2_X)^{½} (1-h^2_Y)^{½} r_E \qquad (9.46)$$

If both traits have low heritabilities, the phenotypic correlation is determined mainly by the environmental covariance, while in the situation with high heritabilities for both traits, the genetic covariance is the more important. A positive genetic correlation may be cancelled out by a negative environmental correlation resulting in a phenotypic correlation close to zero.

Table 9.5. Formula for calculating sums of cross products, and expected mean cross products.

Source of covariation	DF	Sum of cross products	Mean CP	Expected mean cross products
Between sires	s-1	Sire CP	MCPs=Sire SS/(s-1)	$\sigma_{e(XY)}+n\sigma_{d(XY)}+nd\sigma_{s(XY)}$
Within-sires, between dams	s(d-1)	Dam CP	MCPd=Dam SS/(s(d-1))	$\sigma_e^2+n\sigma_d^2$
Within dams	sd(n-1)	Error CP	MCPe=Error SS/(sd(n-1))	σ_e^2
Total	sdn-1	Total CP	Total CP/(sdn-1)	

The resemblance between relatives is utilised when estimating genetic correlations. The phenotypic and genetic covariance between two traits can be estimated in a manner similar to estimation of the phenotypic and genetic variance of one trait with the assumptions that both traits are measured on each individual. The sums of squares and mean sums of squares are replaced by sums of cross-products and mean sums of cross products between traits. Different types of resemblance between relatives, as described

before, may be utilised. Half-sib component of covariance and parent-offspring covariance is most commonly utilised.

By using the same model as in example 9.2.3.2.1, including the random factors sire and dam within sire in addition to the error term, the phenotypic and genetic covariance between two traits measured within the same individual can be estimated from expected mean cross products (CP) according to Table 9.5:

The sire, dam and residual covariance components are equated to the additive genetic, $\sigma_{A(XY)}$, environmental, $\sigma_{E(XY)}$ and phenotypic, $\sigma_{P(XY)}$ covariances as follows:

$\sigma_{A(XY)} = 4\ \sigma_{s(XY)}$

$\sigma_{P(XY)} = \sigma_{e(XY)} + \sigma_{d(XY)} + \sigma_{s(XY)}$

Genetic variances and covariances are combined to estimate the genetic correlation:

$$r_{a(XY)} = \frac{\sigma_{A(XY)}}{\sigma_{A(X)}\ \sigma_{A(Y)}} \qquad (9.48)$$

Estimates of genetic correlations are usually subject to rather large sampling errors as it is composed by three estimated parameters each having an error connected to it. The estimates of genetic correlations are influenced by gene frequencies, and may differ between populations.

Often both traits are not measured on the same individual. Then a method based on multi-trait mixed model must be used including information of relationship between all individuals (Mrode, 1996).

9.6.1.1 Example

From a selection experiment on tilapia, correlations was estimated between harvest weight (HWT) and weight after over wintering (OWT). A multivariate analysis of variance (MANOVA) provided sum of cross products connected to sire, dam within sire and error term between the traits as well as an analysis of variance for each trait. A summary of the output was as following:

Dependent Variable:		OWT	HWT
Source	DF	Mean Square	Mean Square
Model	103	50320.82	28402.46
Error	2764	3518.36	1816.52
Corrected Total	2867		
Source		Mean Square	Mean Square
SIRE		22345.79	12345.36
DAM(SIRE)		15475.16	6788.77
HAPA		104380.54	6960.10
SEX		101332.44	143600.59
POND		12020.33	6.53

Source	Type III Expected Mean Square
SIRE	Var(Error) + 29.426 Var(DAM(SIRE)) + 52.161 Var(SIRE)
DAM (SIRE)	Var(Error) + 29.328 Var(DAM(SIRE))

E = Error SS&CP Matrix			H = Type III SS&CP Matrix for SIRE			H = Type III SS&CP Matrix for DAM(SIRE)		
	OWT	HWT		OWT	HWT		OWT	HWT
OWT	9724741.37	5185219.73	OWT	1184326.78	809084.72	OWT	649956.54	392189.00
HWT	5185219.73	5020853.72	HWT	809084.72	654304.30	HWT	392189.00	285128.18

Source of covariation	DF	Sum of cross products	Mean CP	Expected mean cross products
Between sires	53	809084.72	15265.75	$\sigma_e^2+29.43\ \sigma_d^2 + 52.16\ \sigma_s^2$
Dam(sire)	42	392189.00	9337.83	$\sigma_e^2 + 29.33\ \sigma_d^2$
Within dams	2764	5185219.73	1875.98	σ_e^2
Total	2867			

From the expected mean cross products the sire, dam and error covariance can be calculated. Then genetic, maternal and environmental covariances can be estimated.

$\sigma_{d(XY)} = 254.43$
$\sigma_{s(XY)} = 113.17$
$\sigma_{e(XY)} = 1875.98$

$\sigma_{A(XY)} = 4\sigma_{s(XY)} = 452.68$
$\sigma_{E(XY)} = \sigma_{e(XY)} - 2\sigma_{s(XY)} = 1649.64$
$\sigma_{P(XY)} = \sigma_{d(XY)} + \sigma_{s(XY)} + \sigma_{e(XY)} = 2243.58$

Estimates of variance components for OWT and HWT:

OWT
$\sigma_{d(X)}^2 = 407.69$
$\sigma_{s(X)}^2 = 130.95$
$\sigma_{e(X)}^2 = 3518.36$

$\sigma_{A(X)}^2 = 4\sigma_{s(X)}^2 = 1630.76$
$\sigma_{E(X)}^2 = \sigma_{e(X)}^2 - 2\sigma_{s(X)}^2 = 2702.98$
$\sigma_{P(X)}^2 = \sigma_{d(X)}^2 + \sigma_{s(X)}^2 + \sigma_{e(X)}^2 = 4057.00$

HWT
$\sigma_{d(Y)}^2 = 169.54$
$\sigma_{s(Y)}^2 = 106.21$
$\sigma_{e(Y)}^2 = 1875.98$

$\sigma_{A(Y)}^2 = 4\sigma_{s(Y)}^2 = 424.84$
$\sigma_{E(Y)}^2 = \sigma_{e(Y)}^2 - 2\sigma_{s(Y)}^2 = 1536.90$
$\sigma_{P(Y)}^2 = \sigma_{d(Y)}^2 + \sigma_{s(Y)}^2 + \sigma_{e(Y)}^2 = 2151.73$

Correlations can then be estimated:

$$r_P = \frac{\sigma_{P(XY)}}{\sigma_{P(X)}\ \sigma_{P(Y)}} = 0.76$$

$$r_A = \frac{\sigma_{A(XY)}}{\sigma_{A(X)}\ \sigma_{A(Y)}} = 0.54$$

$$r_E = \frac{\sigma_{E(XY)}}{\sigma_{E(X)}\ \sigma_{E(Y)}} = 0.81$$

9.7 Correlations estimated from multi-trait models based on mixed linear model techniques and pedigree information

Multi-trait models based on mixed linear model techniques where pedigree information is included in the analysis are preferred for estimating genetic covariances between traits. A multi-trait model based on mixed linear model techniques utilises all kinds of relationships between the individuals. This provides the possibility of estimating co-variances between traits, when each trait needs to be described by a different model, or when these traits are not measured on the same individuals. In a multi-trait model the observation vector y can be partitioned into t sub-vectors, i.e:

$$\mathbf{y' = [y_1' \; y_2' \; y_3' \; \ldots\ldots \; y_t']} \tag{9.49}$$

Each vector represents a different trait measured on the same or different individual, and

$$\mathbf{y_i = X_i b_i + Z_i a_i + e_i, \; for \; i=1,2,\ldots..,t.} \tag{9.50}$$

In fish production, measurements may be obtained for growth rate, y_1, and fat content at slaughter, y_2. The model used for one trait could be very different from the model used for the other trait in terms of fixed and random factors included. Some individuals could be measured for both traits, while others may be recorded for only one of these traits. Thus, the complexity of estimating covariance components is greater than for estimating variance components. The fundamental methods remain however unchanged, as multi-trait model techniques are an extension of mixed linear model techniques for single trait models. Consequently, multi-trait models can be used for obtaining BLUE (Best Linear Unbiased Estimator) for the fixed factors, BLUP (Best Linear Unbiased Predictor) for the random factors and estimates for variance components as well as co-variances. All principles and techniques for single trait models apply to multi-trait models. The size of equations is however many times greater than a single trait model increasing the computational difficulty.

The multitrait models can be categorized into three groups:

1 The same model for all t traits, and all individuals are measured for all traits.
2 The same model for all t traits, but some individuals are not measured for all traits.
3 Different models for different traits, and some individuals may not be measured for all traits.

The difficulty of estimating (co)variance components increases from situation 1) to 3). To make the analysis easier, a single trait approach can be used, or transformations can be applied to the data. For situations characterized by group 1) a Single Trait Approach can be used. Variance components could be estimated by a single trait analysis. To estimate the covariance between the traits, new traits are defined as (y_i+y_{i+1}). The variance of the sum of each of two traits is:

$$\mathrm{Var}(y_i+y_{i+1}) = \mathrm{Var}(y_i) + \mathrm{Var}(y_{i+1}) + 2\mathrm{Cov}(y_i,y_{i+1}) \tag{9.51}$$

from where the covariance between two traits can be estimated. The advantage of this method is that the analysis does not require a more complex computation than any single trait analysis. There may however be many analysis to perform.

Canonical Transformation and Cholesky Transformation may be applied to data characterized by group 1) and 2) respectively. Applying a canonical transformation to the data results in a single trait analysis instead of one multi-trait analysis of a giant set of equations. Transformation by Cholesky transformation results in the residual (co)variance matrix of the transformed data becoming an identity matrix making further computations much easier. For further handling of the topic, see Lynch and Walsh (1998).

9.8 Threshold characters

In fish breeding there are many traits of economical importance, which vary phenotypically in a discontinuous manner. Some of these traits are however not inherited in a simple Mendelian manner. They are thought to have a multi-factorial background, which involves a number of physiological functions, influenced by many genes and environmental effects resulting in an underlying continuous variation. The genetic and environmental variation will however not result in a visible phenotypic variation as the phenotype appears in a few distinct classes. Methods for estimating variance components for categorical traits have been developed, based on the concept of a threshold model.

9.8.1 LIABILITY AND THRESHOLD

The value of the underlying trait determines the categorisation of the phenotype. A threshold imposes a discontinuity on the visible expression. As described in Heringstad et al. (2001), the threshold model assumes that phenotypic expression of the categorical trait is determined by an underlying normal distribution, the liability, λ. When the underlying variable is below this threshold level the individual has one form of expression, while when it is above the threshold the individual has the other phenotypic expression, such that the observed binary responses (y) are the result of the following relationship:

$$y_i = \begin{cases} 0 \text{ if } \lambda_i \leq \tau \\ 1 \text{ if } \lambda_i > \tau \end{cases} \qquad (9.52)$$

where τ is a fixed threshold, and y_i=0 and 1 correspond to the phenotypic values. Often liability is assumed to be normally distributed with mean vector μ and (co)variance matrix $\mathbf{R} = \mathbf{I}\sigma_e^2$, where σ_e^2 is the variance of the underlying scale, and **I** is the identity matrix (Searle, 1971). Since these parameters are not identifiable, these parameters are set to some arbitrary values (τ=0 and σ_e^2=1) to denote origin and scale of measurements respectively. Hence it is assumed that the vector of liabilities, given μ, have the distribution: $\lambda|\mu \sim N(\mu, \mathbf{I})$. Then the probability that observation *i* is scored as 1, given model parameter vector, θ, is defined to be:

$$\pi_i = \text{Prob}(Yi = 1|\theta) = \text{Prob}(\lambda_i > 0|\theta) = \phi(\mu_i) \qquad (9.53)$$

where $\phi(.)$ is the standard normal cumulative distribution function.

The assumption of an underlying distribution can be illustrated by the economically important trait litter size. The litter size of cows has only two or three classes. In comparison, fish have large litters, and the trait can be considered a continuous variable because there are large enough classes. There is however no reason to believe that the physiological causes of twinning in cattle are different from those of litter size in fish.

The simplest situation is when we have two phenotypic classes with a single threshold separating them, see Table 5.2. The individuals can only have two phenotypic values, 0 or 1. On a group basis, like family groups, the trait can be expressed in proportion or percentage, referred to as the incidence. The incidence provides a description of the group. The percentage scale is however inappropriate for statistical and genetic analysis as the variance differ according to the mean. The incidence is for this reason converted to mean liability, which is defined as being normally distributed. The unit of liability is then its standard deviation σ. The deviation of the threshold from the mean in standard deviation units of liability, x, relates the mean liability to the incidence, p. Values of x for different incidences are tabulated in Appendix A.

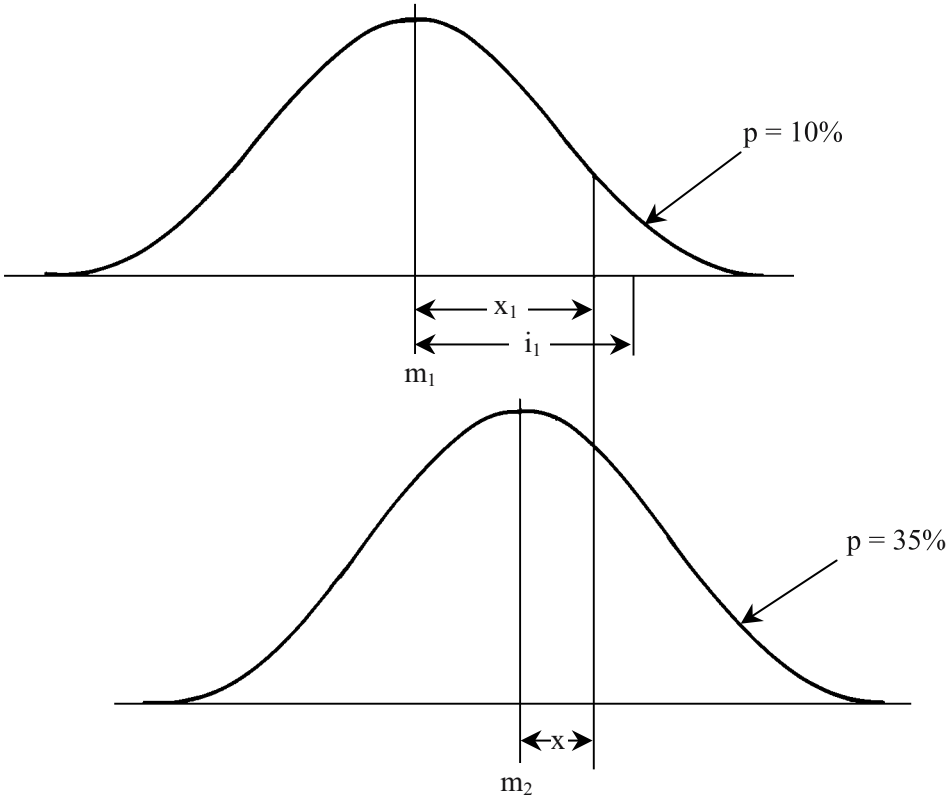


Fig. 9.1 Two populations with different incidences, p, of a threshold character and consequently different liabilities. The variance in liability is the same in the two groups. x is the normal deviate of the threshold from the mean; i is the mean deviation of affected individuals from their group mean.

In Figure 9.1 two groups of individuals with different incidences are show. The thresholds impose the levels where the underlying variables change expression from one to the other phenotypic expression. If the thresholds are defined as being fixed, they are at the same level of liability in all groups), and the groups can be compared.

The group mean is expressed as a deviation in σ units from the threshold. For group 1 in Figure 9.1, 90% of the group express phenotype value '0' and 10% express phenotype value '1', the incidence is p_1=0.10. The threshold is x_1=1.282 σ_1. The mean liability is $m_1 = -1.282\ \sigma_1$. For group 2 in Figure 9.1, 65% of the group express phenotype value '0'. The incidence is $p_1 = 0.35$, the threshold $x_1 = 0.385\ \sigma_1$, and the mean liability is $m_2 = -0.385\ \sigma_1$.

If we can assume that the standard deviation is the same in two groups, the mean liability can be compared. $m_2 - m_1 = -0.385\ \sigma - (-1.282\ \sigma) = 0.897\ \sigma$. i.e. the groups differ by 0.897 standard deviations of liability.

For binary traits, Robertsen and Lerner (1949) presented ways to estimate heritability from a one-way classification model with two sources of variation, among and within family. The heritability on the observed scale with phenotypic expression being 1 or 0, can be estimated by:

$$h^2_{01} = \frac{SSF/p(1-p) - (s-1)}{r(k-s+1)} \qquad (9.54)$$

where p is the proportion above the threshold, SSF is the corrected sum of squares due to families, s is the number of families, r is the additive relationship between family members and:

$$k=\Sigma n_i - \Sigma n_i^2/\Sigma n_i \qquad (9.55)$$

with n_i=number of individuals in the i^{th} family. Heritability can then be estimated on the underlying scale as:

$$h_L^2 = h^2 z^2 /p(1-p) \qquad (9.56)$$

where z is the ordinate of a standard normal density function corresponding to the threshold between '0' and '1'.

Example 9.8.1.1.
In a population of tilapia 1794 fish from 98 full-sib families (s=98, r=0.5) went through a cold challenge test. About 10% survived. The SSF from the variance analysis on the categorical trait was 9.24, while k=1783. The estimated heritability for the categorical trait according to (xx) is h_{01}^2=[9.24/(0.1*0.9) – 98+1]/0.5(1783-98+1)= 0.0067. The estimated heritability of the liability when i=1.76 and z=i*p=0.176, is
$h_L^2 = h_{01}^2 * 0.176^2 / (0.1*0.9)$=0.043.

Falconer (1965) has suggested another method for estimating heritability of liability to diseases from prevalence on relatives. Assumptions of the method are 1) an underlying continuous normal distribution of liabilities in the population where the affected individuals being those whose liabilities exceeds a certain fixed threshold, and 2) the distribution of liabilities in relatives of affected individuals is also normal distributed with variance equal to that in the general population.

If group 1 is the parental generation, while group 2 represent the offspring the heritability can be estimated by the regression of the individual on the mid-parent value expressed by the ratio of response (R) to selection differential (S). The mean of the affected individuals ('1') is $i\sigma$, which is equal to the selection differential, while the response to selection is $m_R - m_P$ where subscripts P and R refer to population and relatives respectively. With incidence of p=10%, $S=i\sigma=1.755\sigma$ in the parent population, the response, R, is 0.897σ, and $h^2=R/S=0.897/1.755=0.51$. This can be applied to any kind of relationship.

First, the correlation of liability between relatives is given by:

$$t = \frac{m_R - m_P}{I} = \frac{x_P - x_R}{I} \tag{9.57}$$

where P=the whole population, R=relatives of the affected individual, and m is the mean as a deviate from the threshold. The heritability is estimated by:

$$h^2=t/r \tag{9.58}$$

where r is the coefficient of relationship between the affected individual and its relatives. If group 1 represent the whole population, while group 2 represent half-sibs of affected individuals, the heritability equals t/(¼).

9.9 Concluding remarks

It must be kept in mind that estimates of genetic and phenotypic parameters are population and time specific. Not surprisingly, genetic constitution differs between populations due to selection, mutations and environmental effects. It explains the fact that estimations vary between studies. This also underlines the need for re-estimation of genetic and phenotypic parameters within a breeding population.

10. BREEDING STRATEGIES

KJERSTI TURID FJALESTAD

10.1 Introduction

The main objective in a breeding program is to move the mean value of the trait in the desired direction for a normally distributed trait, or for a trait with two (or more) discrete classes, such as survival, to increase the frequency of the desired class(es). Changes in the population mean or in class frequencies from one generation to the next are termed selection response. Selection response for a normally distributed trait is illustrated in Figure 7.4.

Selection response or genetic gain can be obtained by applying different breeding strategies. For a long-term breeding objective, the only suitable strategy for a breeding nucleus will be some type of pure-breeding approach for additive genetic improvement. The breeding strategies that may be applied to produce commercial fry are less restricted than breeding within the nucleus. Any kind of crossbreeding, ploidy manipulations and sex manipulation, may be applied if the productivity of the commercial fry can be further improved. Any breeding strategy for commercial fry production that imposes limitation on the progress in additive genetic performance in the nucleus should be avoided (for example the use of highly inbred lines in a crossbreeding program, see below).

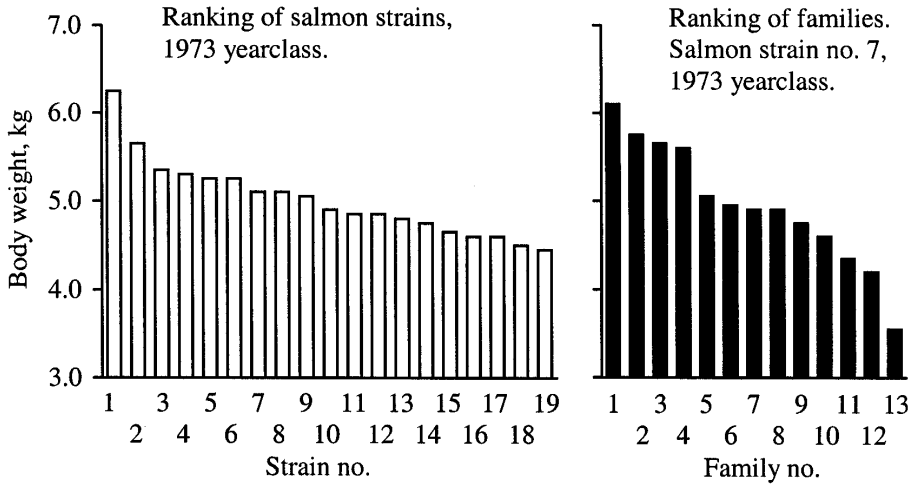


Figure 10.1 Variation in salmon strains and salmon families according to body weight. Reproduced from Gjedrem (1979b) by permission of JUUL Forlag.

T. Gjedrem (ed.), Selection and Breeding Programs in Aquaculture, 145–158.
© 2005 *Springer. Printed in the Netherlands.*

All breeding programs should start with collection, comparison and selection of the best genetic material available (for more details see Chapter 16). The value of testing strains and selecting the best for farming may be equivalent to several generations of within strain selection like shown by Bentsen et al. (1998) for tilapia. Another example of the importance of choice of base population is shown in Figure 10.1, which illustrates the differences in body weight at slaughter between strains of Atlantic salmon and between full-sib families within a strain (Gjedrem, 1979b). The great differences found between strains and between full-sib families within strain state the importance of starting a breeding program with the best available genetic material. The differences between full-sib families within strain also illustrate the possibilities for further improvement through selection.

10.2 Inbreeding

Inbreeding in an infinitely large population is defined as the mating of individuals that are more closely related to each other than individuals mating at random within a population. The populations actually used in most aquaculture programs are finite populations because they possess a limited number of members. All finite populations experience some degree of inbreeding that is based on the number of individuals that contribute progeny to each succeeding generation. Inbreeding in a population is measured by the inbreeding coefficient (F) as defined and described in chapter 6. Inbreeding coefficients express the amount of inbreeding that has accumulated starting from a specific point in the ancestry of the population. The inbreeding coefficient is meaningful only if a specific time in the past is chosen beyond which ancestries will not be considered, and at which time all alleles are considered to be independent.

Knowing the effective population size (N_e) for any breeding structure, the rate of inbreeding per generation (ΔF) can be obtained as:

$$\Delta F = 1/2N_e \tag{10.1}$$

where N_e is a function of the number of sires and dams used as parents for each new generation (other situations influencing the effective population size see chapter 6):

$$N_e=4N_mN_f/(N_m+N_f) \tag{10.2}$$

N_m and N_f are the numbers of males and females, respectively.

Assuming no genetic relationship among sires and dams and when 50 sires and 50 dams are used, N_e = 100 and ΔF = 0.005 or 0.5% per generation. This is likely an acceptable rate of inbreeding for most traits.

The rate of inbreeding depends strongly on the number of the less numerous sexes. If the number of sires is reduced to 30, the number of dams must be increased to 150 in order

to keep $N_e = 100$. And if the number of sires is 20, N_e can never exceed 80, irrespective of the number of dams used.

The general consequence of inbreeding is a redistribution of the genetic variance; the variance component appearing between the means of lines increases while the component appearing within the lines decreases. In other words, inbreeding leads to differentiation between lines and genetic uniformity within lines.

Within a closed population under selection it is impossible to prevent the level of inbreeding to increase over generations. In the Norwegian breeding program for Atlantic salmon and rainbow trout, now run by AquaGen, the increase of inbreeding per generation of selection has been estimated to 1.4% the first 3 generations (Rye and Mao, 1998) and 1.3 % the 6 first generations for rainbow trout (Pante et al., 2001b). Both estimates are based on pedigree informations and Pante et al. (2001b) concludes that pedigree information is necessary for the accurate estimation of rates and levels of inbreeding because effective population size (N_e) provides a poor estimate of the inbreeding rates and level. If the level of inbreeding becomes unacceptable high, some animals from an unrelated population should be used as parents to reduce the inbreeding problem. More information is found in Chapter 6.

10.2.1 APPLICATION OF INBREEDING

- Uniform animals

Inbreeding can be a powerful technique for developing lines for research purposes. Highly inbred lines are genetically stable which is important in connection with the use of "standard" inbred lines for experimental purposes, particularly for laboratory animals to be used in bioassays and other experiments (Komen, 1990).

- Inbred lines combined with crossing

In practical breeding work deliberate inbreeding is of interest only when inbred lines are produced for crossing in order to exploit non-additive genetic variance. Inbreeding in it self is almost universally harmful and the breeder or experimenter normally seeks to avoid it as far as possible.

10. 2.2 INBREEDING DEPRESSION

Inbreeding depression is the effect of inbreeding normally measured as a reduction in the expected performance of the affected trait. Inbreeding depression leads primarily to a reduction of the mean phenotypic value shown by characters connected with reproductive capacity (fecundity, egg size, hatchability) or physiological efficiency (fry deformities, growth rate, survival).

Inbreeding depression is measured as the average performance difference between an inbred population and the base population. Since traits linked to reproduction and physiological efficiency are frequently showing inbreeding depression, it is important to keep the rate of inbreeding at a low level in a breeding program.

10.2.3 RESULTS FROM INBREEDING EXPERIMENTS

Gjerde et al. (1983) studied the effect of three levels of inbreeding (F=0.25, 0.375, and 0.5) on survival and growth rate in rainbow trout, Table 10.1. The average inbreeding depression (over all levels of inbreeding) was 10 % for eyed eggs, 5.3 % for alevins and 11.1 % for fry. No linear relationship between inbreeding and inbreeding depression was found. Growth of fingerlings did not show significant inbreeding depression while growth of adults showed an increasing growth depression with increasing inbreeding. Per 10 % increase in inbreeding coefficient inbreeding depression was found to be 4.5 %, 5.3 %, and 6.1 % at inbreeding levels of 0.25, 0.375 and 0.5, respectively.

Table 10.1 Effect of three levels of inbreeding in rainbow trout on survival of eyed eggs, alevins and fry and on body weight of fingerlings and adult fish. Reproduced from Gjerde et al. (1983) by permission of Elsevier.

	Survival %						Body weight			
Inbreeding	Eyed eggs		Alevins		Fry		Fingerlings		Adults	
coefficient[1]	Control	ID[2]	Control	ID[2]	Control	ID[2]	Control	ID[3]	Control	ID[3]
F = 0.25	95.1	8.9	99.0	2.8	81.4	9.1	12.0	9.8	2.50	11.3
F = 0.375	94.6	15.3	98.4	8.3	72.7	5.6	47.8	22.5	3.15	19.9
F = 0.50	96.8	5.8	96.3	4.8	72.6	18.6	12.2	-4.0	2.96	30.5

[1] Obtained by one (F = 0.25), two (F = 0.375) and three (F = 0.50) successive generations of full-sib mating.
[2] Inbreeding depression (ID = Control - Inbred)
[3] Inbreeding depression ID = $\frac{\text{(Control - Inbred 100\%}}{\text{Control}}$

Su et al. (1996) found also in rainbow trout that the spawning age of females was delayed by 0.53 % per 10 % increase of the inbreeding coefficient and egg number decreased by 6.1%. Inbreeding did not significantly affect egg size or spawning age of males. Inbreeding depression for body weight per 10 % increase in inbreeding varied from 2.26 % to 5.77 % with the tendency that inbreeding depression increased with increasing weight.

In experiments with channel catfish, Bondari and Dunham (1987) reported that inbreeding (25%) increased the number of days required for eggs to hatch but did not significantly influence spawning weight or hatchability.

Typical levels of inbreeding depression are shown in table 10.2. In relation to this it is important to stress that one round of full-sib mating gives an inbreeding coefficient of 0.25 and one round of half-sib mating 0.125, Table 6.1.

As already mentioned deliberate inbreeding is of interest in breeding work only when inbred lines are produced for crossing in order to utilise non-additive genetic variance. In general the resources and time needed to produce, maintain and replace inbred lines will be better utilised by increasing the efforts to improve additive genetic performance of a purebred nucleus (Gjedrem, 1985; Gjerde, 1988).

Table 10.2 Typical levels of inbreeding depression for different traits in fish and shellfish (see also Table 10.1).

Species and trait	Body weight of control	Coefficient of inbreeding	Inbreeding depression per10%ofF	Authors
Rainbow trout:				
Tanks	14	0.25	6.0	Kincaid (1976b)
Tanks	14	0.375	5.4	”
Tanks	203	0.25	9.3	”
Tanks	196	0.375	8.9	”
Tanks	12	0.25	3.9	Gjerde et al. (1983)
Tanks	48	0.375	6.0	”
Tanks	12	0.50	-	”
Cages	2500	0.25	4.5	”
Cages	3200	0.375	5.3	”
Cages	3000	0.50	6.1	”
Eggs hatched, %	59	0.25	1.4	Kincaid (1976b)
Eggs hatched,%	49	0.375	1.3	,, ”
Abnormal fry, %	8	0.25	15	,,
Abnormal fry, %	13	0.37	21.7	,,
Survivalto 147 days%	71	0.25	6.7	,,
Survivalto 147 days%	74	0.375	6.5	,,
Eggs hatched, %	78	0.25	4.2	Aulstad and Kit. (1971)
Abnormal fry, %	-	0.25	4.2	”
Fry survival, %	1	0.25	0.9	”
Atlantic salmon:				
Recapture frequency			17.4	Ryman (1970)
Brook trout:				
Weight at 7 months		0.25	11.1	Cooper (1961)
Weight at19 months		0.25	13.8	“
Coho salmon:				
Tanks	4	0.125	3.4	Solazzi (1977)
Tanks	4	0.25	3.0	“
Carp:				
Growth	500	0.25	6.0	Moav and Wohl. (1963)
Oyster:				
Floating trays	13	0.25	0.9	Longwell (1973)
Floating trays	9	0.25	2.9	Mallet and Haley (1983)
Bags, body weight	94	0.20	8.8	Evans et al. (2004)
Bags, survival %	94	0.20	4.2	”

Breeding approaches to avoid inbreeding fall into three general categories:

- the use of a large random mating population
- the use of systematic line crossing schemes to eliminate the mating of close relatives
- strain crossing to produce hybrids

The use of large random mating populations is the simplest approach and requires only that the breeder take steps to insure that a large number of fish contribute progeny to the next broodstock generation (Bentsen and Olesen, 2002).

10.3 Crossbreeding

Crossbreeding is a well-known method of genetic improvement, which also has applications in aquaculture. Crossbreeding is mating between, species, breeds, populations, strains or inbred lines. The main objective in crossbreeding programs is to exploit non-additive genetic variance (heterosis or hybrid vigour). When lines are inbred without selection the mean of all their crosses is expected to be equal to the mean of the outbred population from which they were derived. Therefore inbreeding followed by crossing cannot produce any improvement, there must be selection at some stage if any improvement is to be made. Crossbreeding should therefore be looked upon as a supplement to a program for additive genetic improvement.

10.3.1 HETEROSIS

Heterosis also called hybrid vigour can be defined as the phenomenon when the offspring surpasses the average of its parents for one or more traits, it is the reverse of inbreeding depression, which is obtained by mating related animals. The two phenomena are almost universally distributed in plants and animals and are particularly associated with reproductive fitness. There are two methods generally used to estimate heterosis. The first one is to compare crossbred progenies with the average of the parental line/strain, the second is to compare the crossbred progenies with the average of the best parental line/strain. If parents arose from different gene pools, crossbreds have increased heterozygosity and therefore a heterosis increment is expected. The extent of the heterosis increment of a given trait depends on the genetic distance between the parent populations.

The relative gains to be made from crossbreeding and selection depend on the magnitude of additive and non-additive variation for the trait or traits in question. If the non-addtive variance is large, substantial gains can be made by crossbreeding (see section 10.3.4).

10.3.2 COMBINING ABILITY

General combining ability (GCA) results from the average effects of a strain (population or line) if it is combined with other strains (populations or lines) in a cross. Any particular cross has an expected value, which is the sum of the general combining abilities of its two parental lines. The cross may deviate from the expectation and this deviation is called the specific combining ability (SCA) of the two lines in combination

GCA = genetic value

SCA = epistatic and dominance deviation, also called "nicking" (this term is often used when the merit of progeny from a mating deviates significantly from the average of the two parents).

The differences of general combining ability are due to the additive variance (A) and A x A interactions in the base population. Differences of specific combining ability are attributable to the non-additive genetic variance and epistasis. Table 10.3 shows how GCA and SCA can be estimated when crossing four populations.

Table 10.3 General and specific combining ability for four populations: A, B, C, and D.

		♂				
		A	B	C	D	
♀	A	-	AxB	AxC	AxD	$S_{1.}$
	B	BxA	-	BxC	BxD	$S_{2.}$
	C	CxA	CxB	-	CxD	$S_{3.}$
	D	DxA	DxB	DxC	-	$S_{4.}$
		$S_{.1}$	$S_{.2}$	$S_{.3}$	$S_{.4}$	S

$s_{.i}$=male cross
$s_{i.}$=female cross
S = total performance for all crosses
S/12 = S/n = mean over all crosses

10.3.2.1 Example
General combining ability (GCA) for population A and B can be estimated as follows:

$$GCA\ (A) = 1/6\ (S_{1.} + S_{.1}) - S/12$$

$$GCA\ (B) = 1/6\ (S_{2.} + S_{.2}) - S/12$$

and specific combining ability (SCA) for A x B and B x A:

$$SCA\ (A \times B) = (A \times B) - GCA\ (A) - GCA\ (B) - S/12$$

$$SCA\ (B \times A) = (B \times A) - GCA\ (A) - GCA\ (B) - S/12$$

It is generally more difficult to measure differences in specific combining ability, and also to utilise these effects in a breeding program. Making and maintaining inbred lines are almost the only means for utilising SCA commercially, although some use can be made of SCA using crossbreeding. The specific combining ability of a cross cannot be measured without making and testing the particular cross.

10.3.3 RECIPROCAL RECURRENT SELECTION

Reciprocal recurrent selection (RRS) is a crossbreeding scheme designed to utilise both general and specific combining ability. The theoretical basis of RRS has been given by Comstock et al. (1949) and Dickerson (1952). RRS begins with two populations, lines A and B. Crosses are made reciprocally, a number of A ♀ being mated to B ♂ and a

number of B ♀ to A ♂. The crossbred progeny are then measured for the character to be improved and the parents are judge from the performance of their progeny. The best parents are selected and the rest discarded, together with all crossbred progeny, which are used only to test the combining ability of the parents. The selected individuals must then be mated again to members of their own line to produce the next generation of parents to be tested. These are crossed again as before and the cycle is repeated. According to Falconer and Mackay (1996) RRS programs are used by commercial breeders of poultry and have given promising results in corn, however, direct comparison with other selection methods has not been encouraging.

The disadvantages of RRS are that the method can be used for multiple spawners only, it requires a complex and extensive testing system and the generation interval will usually be substantially lengthened.

10.3.4 OVERDOMINANCE

When the heterozygote is superior to both homozygotes, the phenomenon is known as overdominance, figure 10.2 (Falconer and Mackay, 1996). Crossing two lines in which different alleles are fixed gives an F_1 in which all individuals are heterozygous, and this is the only way of producing a group of individuals that are all heterozygous. In a non-inbred population no more than 50 % of the individuals can be heterozygous for a particular pair of alleles. Consequently, if heterozygotes of a particular pair of alleles are

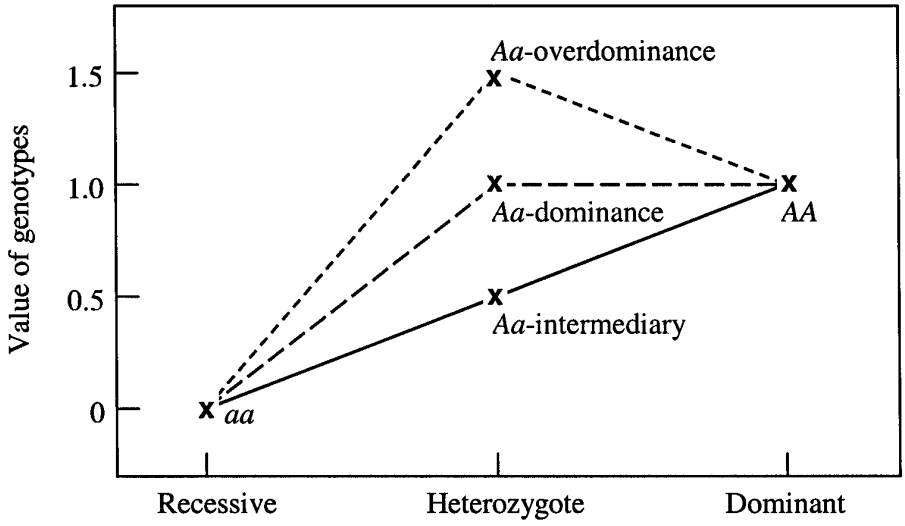


Figure 10.2 Illustration of value of homozygotes and heterozygotes with different effects. Reproduced from Gjedrem (1995a) by permission of JUUL Forlag.

superior in merit to homozygots, inbreeding and crossing will be a better means of improvement than selection without inbreeding.

Furthermore it is only when there is overdominance with respect to the desired character or combinations of characters that inbreeding and crossing can achieve what selection without inbreeding cannot. The existence and importance of overdominance has been much discussed, but the experimental evidence generally suggests the phenomenon of overdominance is not important for most of the characters studied (Falconer and Mackay, 1996).

10.3.5 DIALLEL CROSS

Diallel crossing is a commonly used experimental design for crossing inbred lines or different strains or populations, in which each line/strain/population is crossed with every other line. With p lines this procedure gives rise to a maximum of p^2 combinations. A diallel cross is often used to establish the base population before starting a breeding program. Crossbreeding is frequently used to introduce new genes from strange populations into a local strain. This is usually a simple and very cheap way to improve native strains. However, prior to introduction of a new breed, strain or population should be tested under the existing local conditions.

The base population for the GIFT-project (Genetic Improvement of Farmed Tilapia) was a diallel cross. Four Asian farmed strains and four African wild strains were crossed together in a complete diallel cross (8 x 8 = 64 crosses) to study the magnitude of heterosis in growth performance and survival. The results are shown in Table 10.4 (Bentsen et al., 1998).

Table 10.4 Mean percent heterosis on body weight at harvest across test environments in 56 strain crosses of Nile tilapia. Reproduced from Bentsen et al. (1998) by permission of Elsevier.

	Dam strain							
Sire strain	E2	Gh	Ke	Se	Is	Si	Th	Tw
E2	-	3.8	9.9	11.6	9.3	3.1	9.5	5.7
Gh	4.0	-	-1.5	10.4	2.6	3.9	14.0	3.7
Ke	10.6	-1.5	-	-0.8	4.8	1.9	5.6	1.6
Se	12.3	10.3	-0.8	-	6.3	0.3	2.8	0.0
Is	10.4	2.8	4.9	6.7	-	0.6	7.0	0.4
Si	2.8	3.5	1.7	0.2	0.4	-	5.8	2.5
Th	10.0	13.7	5.4	2.7	6.5	6.0	-	-5.2
Tw	6.0	3.8	1.7	0.1	0.3	2.8	5.8	-

10.3.6 THREE-WAY AND FOUR-WAY CROSSES, BACKCROSSES

In a three-way cross the F_1 of two lines, in which for example high productivity is required is crossed with a third line. In a four-way cross two F_1's of different pairs of lines are crossed. Using backcross two lines only are involved, the F_1 being mated to one of the lines used in the first cross. Crossing is widely used in animal production most of the animals produced for meat being the progeny of either a three-way cross or a backcross. In aquaculture these methods are rarely used.

10.3.7 SYNTHETIC POPULATIONS

A diallel cross to test the potential populations is usually the starting point for making a synthetic population like done for Atlantic salmon, rainbow trout and tilapia. (Gunnes and Gjedrem, 1978; Gjedrem et al., 1987; Refstie, 1990; Bentsen et al., 1998).

Synthetic populations are built from different number of parental populations, breeds, stocks or lines. When making a synthetic population the breeders try to create something new which combines the advantages of the parental populations. Crossing a number of selected inbred lines or different populations and allowing the F1 and later generations to mate at random or more usually with planned mating creates this new population.

Synthetic populations are expected to harbour more heterozygosity than the parental strains and they should show some heterosis gain. The heterosis may, and often will, be reduced by inbreeding subsequent to a reduction in the synthetic population size. Also if losses due to recombination are important, these will be evident in later generations of the synthetic population.

10.3.8 APPLICATION OF CROSSBREEDING IN FISH AND SHELLFISH

The first step in determining if crossbreeding has a place in the breeding strategy for a particular species is to evaluate all possible crosses between different strains or species for the economic traits in question. If the number of strains available is large, one must select the crosses that are most likely to give valuable results. It may be advantageous to make use of strains with very different origin and also to use strains that, in combination, have favourable traits. In Israel crossbreeding programs are currently in operation, which use strain crosses of common carp (Wohlfarth et al., 1983).

Secondly inbred lines should be developed and the crosses tested under natural conditions to find the most valuable cross for farming. This breeding system aims particularly at utilising non-additive genetic variances. One of the practical difficulties here would be to keep the inbred lines running because of high mortality (inbreeding depression). Bakos (1979; 1987) reported results where inbred lines of common carp were used in a crossbreeding program.

Third and last, if possible a reciprocal recurrent selection (RRS) program should be evaluated, to determine the relative importance of general and specific combining

abilities. RRS can only be used for multiple spawners and can therefore not be applied in for example Pacific salmon. It would also be very difficult to practice in Atlantic salmon since the majority of males die after the first spawning together with a high number of females. Other species, such as tilapia and rainbow trout, may be more suitable candidates for the application of RRS schemes.

One significant advantage, using crossing between selected lines in a breeding program is that this enables the breeders to protect their genetic improved material. By selling crossbred animals only, the pure-bred seedstock are not released.

10.3.9 EXAMPLES AND RESULTS FROM CROSSBREEDING EXPERIMENTS

Chevassus (1979) reviewed the status of hybridisation between species of salmonids. He concluded that in most cases the hybrids farmed in the same environment, as the parental species show intermediary or, at best, equal growth to that of the better of the parents. This is in accordance with results from Refstie (1983a) after hybridising of four salmonid species, Atlantic salmon, brown trout, sea trout and Arctic char. Neither growth nor survival of hybrids did exceed the performance results of Atlantic salmon. In some experiments, better results have been found for survival rate, the hybrid often being similar or even superior to the hardiest species of hybrid combinations (Chevassus 1979).

Table 10.5 Average body weight at slaughter (kg) and survival (%) in the sea rearing period of crosses among five Norwegian strains of Atlantic salmon. In parenthesis average performance of pure bred strains. Reproduced from Gjerde and Refstie (1984) by permission of Elsevier.

	Body weight			Survival		
Strain	Average of reciprocal crosses	Average of purebred strains	Heterosis[1]	Average of resiprocal crossees	Average of purebred strains	Heterosis[2]
Rana x Nidelva	3.90	4.17	-6.5	50.9	44.6	6.3
Rana x Surna	3.64	3.73	-2.4	51.3	57.6	-6.3
Rana x Driva	3.85	4.23	-9.0	52.1	55.0	-2.9
Rana x Eidfjord	4.34	4.10	5.9	63.7	46.3	17.4
Nidelva x Surna	3.45	3.42	0.9	47.0	44.9	2.1
Nidelva x Driva	4.10	3.92	4.6	40.5	42.3	-1.8
Nidelva x Eidfjor	1.01	3.92	2.3	38.3	33.7	4.6
Surna x Driva	3.33	3.48	-4.3	36.7	55.3	-18.6
Surna x Eidfjord	3.56	3.47	2.6	53.0	46.7	6.3
Driva x Eidfjord	3.66	3.98	-8.0	49.1	44.1	5.0
Average	3.78	3.84	-1.6	48.3	47.1	1.2

Body weight and survival for purebred, respectively: Rana 4.48, 57.2; Nidelva 3.86, 31.9; Surna 2.97, 57.9; Driva 3.98, 52.7 and Eidfjord 3.97, 35.4.
[1](Crosses – Purebred) x 100/ Purebred
[2] Crosses – Purebred

Benzie et al., (1995) found no indication of hybrid vigour in growths rates when crossing two species of tiger shrimp, *Penaeus monodon* and *P. esculentus*. Growth rates of the hybrids were similar or lower to that of pure *P .monodon.*

Gjerde and Refstie (1984) investigated the heterosis effect between crosses of five Norwegian strains of Atlantic salmon. They did not find a significant heterosis effect for either growth rate or survival rate (Table 10.5). Likewise, Friars et al. (1979) found no heterosis effect for growth rate of Atlantic salmon fry. However, in rainbow trout Gall (1975) and Ayles and Baker (1983) reported significant heterosis for body weight among crosses of rainbow trout strains.

Systematic crossbreeding has been performed among common carp varieties. The significance of heterosis for growth rate, survival and cold tolerance among crosses of wild and domesticated European, Russian, Chinese and Japanese strains have been repeatedly reported (Table 10.6) (reviewed by Hulata, 1995). Wohlfarth (1993) summarised experimental data collected over more than two decades of research in Israel on common carp and concluded "heterosis for growth is a common but not universal phenomenon in carp." As a rule, heterosis was not found when one of the parental lines was Dor-70, a line generated in a long-term mass selection experiment for faster growth. Gjerde et al. (1999) estimated heterosis for body weight and survival in rohu carp (*Labeo rohita*) and concluded that cross-breeding of Indian stocks of rohu carp seems to have little practical significance.

Table 10.6 Crossing between different races of common carp. Growth rate, g. Reproduced from Moav et al. (1975) by permission of Blackwell Publishing Ltd.

Crosses	Average of crosses	Average of pure breeding	Difference crosses-purebreds	Difference crosses-best parent
Big-Belly (336) x Nas (439)	491	388	103	52
x Gold (443)	470	390	80	27
x Dor (556)	500	446	54	-56
Nas x Gold	559	441	118	116
x Dor	558	498	61	2
Gold x Dor	516	500	17	-40
Average	516	444	72	17

The GIFT crossbreeding experiment with tilapia (Bentsen et al., 1998), already mentioned, showed that only seven out of the 22 crosses that expressed a significant heterosis were better performing than the best pure strain and the largest gain was about 11 %. In general non-additive genetic effects on body weight were modest compared to the additive plus reciprocal effects.

Both Wohlfarth (1993) and Bentsen et al. (1998) have reported results that indicate that the expression of non-additive genetic effects may be more sensitive to environmental variation than additive effects. Heterosis may then be poorly expressed in some farm

environments because of genotype by environment interactions affecting the non-additive genetic performance. In the case specialised hybrids may have to be produced for certain farm environments.

Knibb et al. (1997) found that crossbreeding among seabream strains yielded little heterosis, this was explained by lack of inbreeding and genetic differentiation. Reviewing results from several crossbreeding experiments Knibb (2000) concluded that over all species, hybrids usually resemble the average of their parents. Given the large number of attempts to produce new fish hybrids, remarkably few (apparently less than 1 %) have resulted in sustained commercial production.

The search for infertile hybrids, which do not divert food into gonads and therefore have superior production traits, may become very important. As unisexual off-spring have been obtained in several interspecific crosses in tilapia such monosex (male) cultures are recognised as the best solution to over-population caused by the high fecundity of tilapia under almost any pond condition. In addition, males also grow faster than females. Pruginin et al. (1975) list several crosses which have given 100 percent male offspring while Hulata et al. (1983) propose to use progeny testing to ensure getting 100 percent male progeny in tilapia. A promising cross between *Oreochromis niloticus* and *O. aureus* tilapia in Israel is reported to produce nearly all male offspring (Hulata, 1995).

10.4 Purebreeding

The breeding method or strategy for additive genetic improvement within a population is known as purebreeding, and the method to choose for continuous genetic improvement over a long period of time. Mating of close relatives must be avoided and individuals that possess a majority of positive (desirable) genes are selected as parents for the next generation. Individuals that possess a majority of positive gene alleles normally show good production results. These "good genes" and properties are transferred to their offspring. Individuals that possess a majority of positive genes are said to have a high breeding value.

The breeding value of an individual cannot be measured directly. Neither can it be measured with 100% accuracy. The true breeding value will therefore remain unknown and to a greater and lesser extent be masked by systematic and stochastic environmental effects and also by effects caused by interactions among the genes, see chapter 14 for discussion of genotype-environment interaction.

The breeding value can primarily be estimated by recording the product of the genes, i.e. the phenotypic values of the trait(s) (or the use of genetic markers linked to QTL as described in Chapter 19). The phenotypic records may be obtained from the individual itself or from relatives as full- and half-sibs, progeny or parents. Records on relatives can be used because the individual and its relatives share common genes. Information from close relatives is in general more valuable than information from distant relatives. Records on full-sibs are thus more valuable than records on half-sibs because the

individual shares a larger proportion of common genes with its full-sibs than with its half-sibs. Records on progeny are of particular interest as the breeding value of an individual is strictly defined as the value of an individual judged by the mean value if its progeny.

11 SELECTION METHODS

KJERSTI TURID FJALESTAD

11.1 Introduction

The aim of selection is to identify and select as parents for the next generation the individuals whose progeny, as a group, have the highest possible additive genetic merit for the trait or traits in question. This is almost equivalent to selecting as parents those whose own additive genetic merit is highest. The basic effect of selection is to change the gene frequencies, and the effects that can be observed are changes of the population mean. If there are no differences of fertility among the selected parental individuals or of viability among their progeny, then the gene frequencies are the same in the offspring generation as in the selected parents.

Available selection methods to fish breeders include individual selection, pedigree selection, family selection, within family selection, combined family and within family selection, and progeny testing. The efficiency of each method can be predicted by calculating the expected genetic response for a given set of parameters. The efficiency of selection is partly dependent on how accurately the breeding values of individual animals are evaluated (see section 11.8).

In this chapter we distinguish between six different methods of selection of breeding animals. For each method, we describe from which individuals or relatives we obtain our information for the estimation of breeding values of selection candidates. The objective for all methods is to maximise the probability of correctly ranking of the potential breeding animals (on true additive genetic merit). This is equivalent to maximising the correlation between the true and estimated breeding value. This correlation (r_{TI}) is frequently termed the accuracy of the breeding values and is an important parameter as it is directly proportional to the expected response to selection.

What selection method to choose is dependent on several factors including the heritability of the trait(s), the nature of the trait, recording methods and the reproductive capacity of the species. A description of each method is given in the following.

11.2 Individual selection

Selection based on an individual's own performance or phenotype is called individual or mass selection. This is a well-known and widely used method of selection in animal breeding, and for most aquaculture species the only method practised. Individual selection is usually the simplest method to operate and in many circumstances it yields the most rapid response. However mass selection may be unsuitable if there are large uncontrolled systematic environmental variation (for example age differences). Additionally, there is no control of inbreeding with mass selection, and this has caused

T. Gjedrem (ed.), Selection and Breeding Programs in Aquaculture, 159–171.
© 2005 *Springer. Printed in the Netherlands.*

problems in a number of fish breeding programs (Moav and Wohlfahrt, 1973; 1976; Hulata et al., 1986; Teichert-Coddington and Smitherman, 1988).

A prerequisite for using individual selection is that the trait(s) can be measured on the breeding individual itself while being alive. The method is consequently difficult to practice for meat quality traits, disease resistance and for age at sexual maturity when the frequency of maturing or immature fish is low (<10-15%) and it is not efficient on traits with low heritability. The accuracy of the selection method is equal to the square root of the heritability of the trait (when h^2=0.25 accuracy r_{TI} = 0.50).

In fish individual selection is of particular interest for growth rate as the heritability for this trait is fairly high (h^2=0.20-0.40) in most species. Since there are many advantages with individual selection, methods should be developed for recording also other traits than growth rate like for example ultrasound of live fish to assess fat coverage.

When applying individual selection it is of vital importance that the environmental influence is kept the same for all individuals that are to be compared at any stage of the life cycle. In this manner, the probability of correct ranking of the individuals is maximised. Differences between individuals or groups of individuals for environmental factors like water temperature and salinity, different tanks, ponds or cages, density, light condition, and type of food and feeding regimes, may reduce the accuracy of the selection substantially and in that way reduce the possibility for genetic improvement. To obtain as equal environmental conditions as possible all individuals to be compared should be hatched at the same day or within a few days period and thereafter reared under identical environmental conditions.

If we want to select individuals reared in different environments, for example in different tanks, or in different cages on the same or different farms; a particular precaution is needed. Due to environmental differences, absolute trait values are not directly comparable. Instead, the mean value for each environmental group should be calculated and the individuals ranked and selected based on their deviations from their group mean.

If the environmental groups are significantly different in their mean values, selection based on deviations will result in more individuals being selected from the group having the highest mean value. This bias is a result of the dependency between the mean and the variance for many traits and frequently observable for body weight. Comparable deviations can be obtained by first multiplying all observations in an environmental group with a constant that is equal to the ratio between the mean value for the group chosen as the base and the mean value for the actual group (for more information see Chapter 5).

When individual selection is the only selection method practised, it is not usual to tag individuals and in that way keep track of the genetic relationship among them. Neither is it possible from a practical point of view to keep many full and half-sib groups separate until time of selection. Therefore, to avoid mating of close relatives and rapid

accumulation of inbreeding, the number of breeding animals should be kept fairly high, at least 50 mating pairs per generation (Bentsen and Olesen, 2002) thereby securing a large effective population size.

11.3 Pedigree selection

By this selection method, breeding animals are selected based on their parents, grandparents or more remote ancestor's performances or breeding values. However, an individual receives a random sample of half of its chromosomes or genes from each parent, opening a vast number of possible new chromosomes or gene combinations among the offspring. This segregation of genes for each new generation may result in substantial deviation in breeding values among offspring.

Selection based on performance of ancestors is generally of limited value if other information is available. The weakness of pedigree information in improving the accuracy of selection lies in the fact that information is generally available on only a few ancestors and the effects of genetic segregation become large after one or two generations.

The accuracy of this selection method will therefore not be high. As a result pedigree selection is not much used as the only method of selection in modern animal breeding.

The accuracy of a breeding value based on both parents' phenotype is equal to $r_{TI} = \sqrt{1/2h^2}$ and consequently $r_{TI} = 0.35$ for $h^2 = 0.25$, where h^2 is the heritability of the trait in question.

11.4 Family selection

Family selection refers to a selection method in which family groups are ranked according to the mean performance of each family and whole families are saved or discarded (Lush, 1947). The individuals saved as breeders for the next generation are either all the individuals in selected families or randomly chosen individuals taken equally from all selected families. Because selection is among whole families, the selection differential is a function of differences among families, not differences among individuals.

Individual values are not used except in so far as they determine the family mean. The families may be of full-sibs or half-sibs, families of more remote relationships being of little practical significance.

The efficiency of family selection rests on the fact that the environmental deviations of the individuals tend to cancel each other out in the mean value of the family. Accordingly the phenotypic mean of the family comes close to being a measure of its genotypic mean. The advantage of family selection over other types of selection is

greater when environmental deviations constitute a large part of the phenotypic variance. The chief circumstance under which family selection is to be preferred is when the trait selected has a low heritability. With low heritability the use of family mean give an increased accuracy when estimating breeding value. On the other hand, environmental variation common to members of a family impairs the efficiency of family selection. If this component is large it will tend to swamp the genetic differences between families and family selection will be correspondingly ineffective.

To reduce the common environmental component to a minimum, the environment for all families should be standardised as far as possible in the period the families are kept separate. In addition, individuals from all families should be tagged as early as possible and thereafter reared together in the same tank, pond or cage, communal rearing. Refstie and Steine (1978) estimated the common environmental component in Atlantic salmon, due to rearing families in separate tanks for a period of 6 months, and found that this component contributed about 10 % to the total variation in growth rate. If the common environmental component is large, splitting of families in replicate/triplicate tanks should be considered.

Another important factor affecting the efficiency of family selection is the number of individuals in the families; the family size. The larger the family, the closer is the correspondence between mean phenotypic value and the mean genotypic value. The high reproductive capacity in aquatic animals makes family selection important for these species.

Summarising to this point, the conditions that favour family selection compared to individual selection are low heritability, little variation due to common environment and large families.

Table 11.1 Expected responses to selection based on family for survival at different mean survival rates after one generation. A heritability of 0.1 on the underlying liability scale and a family size of 10 full-sibs and 20 half-sibs is assumed.

Survival rate (%)	Response (%)	Relative response
50	5.04	100
60	4.84	96
70	4.24	84
80	3.22	64
90	1.77	35

Another great advantage with family selection is that, based on phenotypic observations on full-sib and/or half-sibs, breeding values can be estimated for traits that cannot be measured on the individuals that are to be used as parents. Carcass quality traits and

disease resistance can therefore be included in the breeding objective by applying family selection. Family selection is also far more effective than individual selection for threshold traits such as age at sexual maturity, particularly at low frequencies or incidence of the trait. An incidence of 50 per cent gives the best discrimination between families as shown by an example for survival in Table 11.1.

In order to keep the rate of inbreeding low and the intensity of selection high, the number of family groups bred and measured should not be smaller than 50-100. In the period prior to tagging, the family groups have to be kept separately. This makes family selection costly in terms of space. If breeding space is limited in this period, the intensity of selection that can be achieved under family selection may be quite small.

The accuracy of family selection is dependent on several parameters; the heritability of the trait (h^2), the family size (n) and family type (full and/or half-sibs), and the variation due to common environment (c^2). Accuracies (r_{TI}) for some parameter combinations are given in Table 11.2.

Table 11.2. Correlation between true and estimated breeding value (r_{TI}) for a normally distributed trait recorded on n full-sibs and n(m-1) half-sibs and when the environmental variance common to full-sibs (c^2) is 0% and 10 % of the total variance (figures in parentheses are the correlation when the trait is also recorded on the breeding candidate itself).

h^2	No.of full-sib groups (m)	Number of fish per full-sib group (n)			
		$c^2 = 0.00$		$c^2 = 0.10$	
0.10	1	0.42(0.48)	0.60(0.64)	0.33(0.42)	0.39(0.47)
	2	0.44(0.50)	0.61(0.64)	0.35(0.44)	0.41(0.49)
	3	0.45(0.51)	0.61(0.64)	0.37(0.45)	0.43(0.50)
	4	0.47(0.52)	0.62(0.65)	0.39(0.46)	0.44(0.51)
	5	0.48(0.53)	0.62(0.65)	0.40(0.47)	0.45(0.52)
	10	0.50(0.55)	0.62(0.65)	0.44(0.50)	0.48(0.53)
0.25	1	0.54(0.65)	0.66(0.73)	0.45(0.60)	0.51(0.64)
	2	0.56(0.66)	0.66(0.73)	0.47(0.61)	0.52(0.64)
	3	0.56(0.66)	0.67(0.73)	0.49(0.62)	0.54(0.64)
	4	0.57(0.66)	0.67(0.73)	0.50(0.62)	0.54(0.65)
	5	0.57(0.67)	0.67(0.73)	0.51(0.63)	0.55(0.66)
	10	0.58(0.67)	0.67(0.73)	0.53(0.64)	0.56(0.66)

With a family structure of full-sibs only, the upper limit of the accuracy is $r_{TI} = 0.71$, which is the square root of the additive genetic relationship among full-sibs, assuming the environmental variance common to full-sibs is equal to zero. The upper limit of the accuracy for half-sibs is $r_{TI} = 0.50$.

There are two major limitations with family selection. Intensive family selection can quickly result in rapid accumulation of inbreeding because whole families are selected. Another weak point of family selection is that as only 50 % of the additive genetic variation is expressed between families, only 50 % of the variation can be utilised. The other 50 % of the additive genetic variation are expressed within families.

11.5 Within family selection

The criterion of within family selection is the deviation of each individual from the mean value of the family to which it belongs. This type of selection is the reverse of family selection, the family means being given zero weight in the selection decision. The chief condition under which this method has an advantage over the others is when there is a

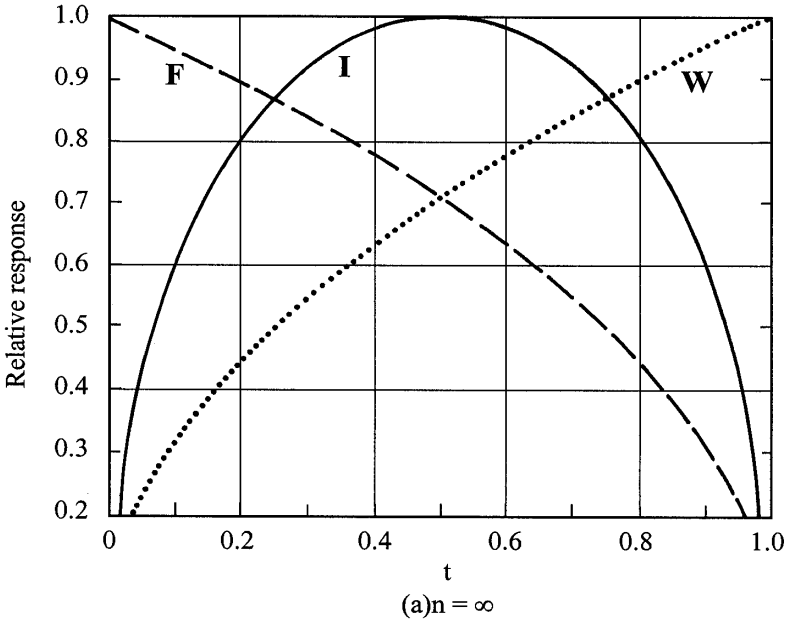


Figure 11.1. Relative merits of the different methods of selection. Responses relative to that of combined selection plotted against the phenotypic intraclass correlation, t. I = individual selection; F = family selection; W = within family selection. Reproduced from Falconer and MacKay (1996) by permission of Pearson Education Inc.

large component of environmental variance common to members of a family (Figure 11.1). Selection within families would eliminate this large non-genetic component from the variation operated on by selection (Uraiwan and Doyle, 1986).

An important practical advantage of selection within families, especially in laboratory experiments, is that it economises breeding space, unlike family selection. For selection

within full-sib families the accuracy, for h^2=0.25, is $r_{TI} = \sqrt{1/2h^2}$ and r_{TI} = 0.35. This accuracy is equal to a breeding value based on both parents' phenotype (see section 11.3). Within family selection has low efficiency compared to most other selection methods (Gall and Huang, 1988 a, b).

11.6 Progeny testing

Progeny testing offers the most direct assessment of an individual's breeding value because individuals are selected on the basis of the performance of their offspring. This method of selection is widely applied in breeding programmes of less prolific species like for example dairy and beef cattle, sheep and goats. In prolific species like fish it is of much less advantageous since family selection can be applied.

The criteria of selection, as the name implies, is the mean value of an individual's progeny. The progeny families are usually half-sibs. The selection is made among the parents on the basis of their progeny means. Progeny testing suffers from the serious drawback of a much lengthened generation interval, because the selection of the parents cannot be carried out until the offspring have been measured. Frequently the generation interval will be doubled. This limitation makes progeny testing impossible or difficult in species, which either spawn only once or suffer high mortality during or after spawning. This is true for Pacific salmon where all fish die after spawning and in Atlantic salmon where a considerable percentage of the animals die after spawning. Deep-freezing of semen is possible in salmonids (Stoss and Refstie, 1983), but is not sufficiently reliable to be recommended used in practice.

Interestingly, progeny testing is the only selection method for which the accuracy can be 100% (Gjerde, 1991) with large progeny groups and irrespective of the size of the heritability.

11.7 Combined selection

The term, combined selection is used simply to denote that more than one method of selection is used in a breeding plan. This method, ideally, combines in an optimal way all available sources of information that can add to our knowledge about the breeding value of an animal, information recorded on the animal itself, information about full-sibs and/or half-sibs and progenies as well as pedigree information. It represents the general solution for obtaining the maximum rate of genetic gain, and the other simpler methods are special cases of this method. Combined selection is therefore in principle always the best method.

The simpler methods usually combined are family and individual selection, family deviations are taken into account in addition to the mean phenotypic value of individuals. Accuracies (r_{TI}) for some parameter combinations are given in Table 11.2.

Selection indexes are the most efficient method to combine information from an individual and its relatives as well as information from several traits. All information is combined into an index of merit where the traits are weighted according to their relative economic value. This is described in detail in Chapter 13.

11.8 Prediction/Expected response

Selection response, measured as incremental improvement per year, is the most critical aspect of the efficiency of a breeding plan. The length of time required to complete a generation, the generation interval, can be excessive for some selection methods. When this occurs, a good response per generation may not be practical due to the number of years required to achieve the response. It is also difficult to discuss selection methods because the procedures involved can have a major effect on selection intensity. To overcome this difficulty, selection responses for various selection methods are standardised by calculating responses as a ratio using responses to individual selection as a common denominator. One other practical consideration is the influence of various selection methods on rate of inbreeding.

Genetic gain depends on:

- how well the animals are evaluated, or the accuracy of prediction
- the amount of selection or selection intensity
- the magnitude of the genetic differences among animals or standard deviation of additive genetic values and
- how rapidly better younger animals replace their parents, known as generation interval

The expected genetic gain (ΔG) or response to selection per year is dependent on four parameters and was given in equation (7.4):

$$\Delta G = \frac{i \times h^2 \times \sigma_p}{L}$$

where:

i = the standardised selection differential, also called the intensity of selection. Because of the high fertility in most aquatic animals a very high intensity of selection can be applied (Appendix A)

h^2 = the heritability of the trait (some examples/estimates are given in Table 5.7)

σ_p = the phenotypic standard deviation; i.e., the square root of the phenotypic variance (σ^2_p)

L = the generation interval is defined as the average age of the parents at the birth of their selected offspring. This parameter will vary from one species to another and according to management practice. For Atlantic salmon produced from 1-year old smolt, the generation interval will be 4 years when the broodstock is selected after two winters in the sea. For rainbow trout it can be as short as 2 years, but frequently 3 years. It is important to keep the generation interval

short to expedite the selection progress. Information about length of generation interval is given in Table 8.1.

A more general formula, applicable to all methods of selection is:

$$\Delta G = \frac{i \times r_{TI} \times \sigma_G}{L} \qquad (11.1)$$

Where i and L is as described above and

r_{TI} = the accuracy of selection; i.e. the correlation between the true and estimated breeding value.

σ_G = the genetic standard deviation; i.e., the square root of the additive genetic variance (V_G).

The expected response to selection is therefore directly proportional to the size of the accuracy of selection. The efficiency of different methods of selection can therefore be measured as the ratio between their accuracy of selection assuming i, σ_G and L the same for all methods.

11.9 Correlated response

Natural selection will enable animals to adjust to their environmental conditions whether they are in the wild or under crowded farming conditions. Natural selection is therefore important for domestication. Artificial selection for economic important traits such as growth rate and survival will facilitate the rate of domestication (Doyle, 1983). The result will probably be animals that adapt better to life in captivity with lower levels of stress. Farmed Atlantic salmon in Norway have undergone seven generations of selection for increased growth rate. The fish in the later generations seems to be less sensitive to environmental stress than genetically wild fish. These changes are difficult to measure, but are easily seen.

Correlated response to artificial selection depends on the genetic correlation between the traits, their heritability, the selection differential and the phenotypic variance for the trait not directly selected for. The correlated repose in trait P_2 (CR_2) when selection is for trait P_1can be predicted by equation (7.6):

$$CR_{P2} = i.h_{P1}.h_{P2}.r_G.\sigma_{P2}$$

where i is the intensity of selection, h^2_x is the heritability of trait x, r_G is the genetic correlation and σ_{P2} is the phenotypic standard deviation of the correlated trait.

Two causes of genetic correlation between characters are described in breeding theory; linkage and pleiotropy. Linkage is the situation in which different loci are situated close together on the same chromosome preventing the genes from segregating independently at

meiosis. Pleiotropy is simply the situation where a single gene affects two or more traits. The actual genetic correlation between two traits is the net effect of pleiotropy and linkage (Falconer and Mackay, 1996).

In livestock species negative side effects of selection for high production efficiency seems to be more at risk for behavioural, physiological and immunological problems (reviewed by Rauw et al., 1998). Beilharz et al. (1993) explained these side effects by the resource allocation theory, which states that fitness is a product of many major fitness components. Consequently when planning and implementing selective breeding in aquatic species precautions to avoid negative side effects of selection should be taken. Careful monitoring of possible correlated traits is needed and more basic knowledge of welfare and behavioural needs.

11.10 Indirect selection

In a selection program it may be desirable to include correlated traits with no economic value in order to improve a trait through its correlated response rather than select for it directly (Gjedrem, 1967). If we want to improve character A, we might select for another character B, and achieve progress through the correlated response of character A. This is indirect selection; selection applied to some character other than the one it is desired to improve. Indirect selection can only be expected to be better than direct selection if the correlated trait has a substantially higher heritability and the genetic correlation is high (Falconer and Mackay, 1996). However, practical considerations may make indirect selection preferable.

There are three practical considerations that may make indirect selection preferable:

- If the desired character is difficult to measure with precision, the errors of measurement may so reduce the heritability that indirect selection becomes advantageous. This is frequently true for disease resistance

- If the desired character is measurable in one sex only, but the secondary character is measurable in both, a higher intensity of selection will be possible by indirect selection

- The desired character may be costly to measure as for example the efficiency of feed conversion. Then it may be economically better to select for an easily measured correlated character, such as growth rate

In fish disease resistance is a candidate trait for indirect selection since disease resistance is difficult to measure on individuals. Today this can be done by including survival results after challenge tests in the breeding goal or survival results during grow out. One major drawback for both methods is that it can be done on a family basis only. By using an indirect measure correlated to disease resistance this could be done on an individual basis. Immune parameters and physiological parameters such as lysozyme (Røed et al., 1992; 1993) haemolytic activity (Røed et al., 1990; 1992; 1993) antibody response to specific

antigens (Eide et al., 1994; Strømsheim et al., 1994a, b; Lund et al., 1995; Fjalestad et al., 1996), $\alpha_2$makroglobulin, antiplasmin (Salte et al., 1993) and cortisol (Refstie, 1982; Fevolden et al., 1994) have shown genetic variation and are examples of possible traits for selection for increased disease resistance in Atlantic salmon. They can be all recorded on the breeding candidates and their full and half-sibs. Since blood samples can be taken prior to selection of broodstock the generation interval will not be increased.

Table 11.3. Relative expected response to direct selection for survival at an overall survival of 50% and 90% after one generation. A family size of 10 full-sibs and 20 half-sibs and an environmental correlation of zero are assumed. Reproduced from Fjalestad et al. (1993) by permission of Elsevier.

			50% survival		90% survival	
	h^{2*}	r_G	Family selection	Individual + family selection	Family selection	Individual + family selection
Direct selection:						
Survival	.1	-	100	-	100	-
Indirect selection:						
Correlated trait I	.3	.3	44	52	55	66
Correlated trait II	.3	.5	73	87	92	110
Correlated trait III	.3	.7	103	122	129	154
Correlated trait IV	.6	.7	113	145	142	183
Combine correct and indirect selection:						
Survival + corr. trait I	.3	.3	105	108	109	115
Survival + corr. trait II	.3	.7	125	140	146	167

* Estimate on the assumed underlying liability scale

Indirect selection for survival based on a correlated trait with no economic value results in a much lower response in survival compared to direct selection (Table 11.3). This is the expected correlated response in survival when information of a correlated trait is included in the selection index and survival is the only trait included in breeding goal. The genetic correlation between survival and the correlated trait must be large if indirect selection shall compete with direct selection. Table 11.3 also shows that more can be gained by including information on correlated traits when overall survival rates are high, in particular if indirect records on the breeding candidate itself are included. The same would apply to low survival rates.

By combining direct and indirect selection for survival (at an overall survival of 50 %) the expected response increases by 5 % (family selection) or 8 % (combined family and individual selection) when the correlated trait has a heritability of 0.3 and a genetic correlation of 0.3 to survival (Table 11.3). These genetic parameters are close to e.g. estimates obtained for growth rate. If the genetic correlation between the traits is high (0.7) the response will increase considerably. So the most effective use that can be made of a correlated character is in combination with the desired trait, as an additional source of information about the breeding values of individuals.

Table 11.4 Values of the correlation between the additive genetic merit of the individual being evaluated and the phenotypic measure used as basis for selection (r_{TI}) at different heritability levels.

Information	Heritability, h^2				
	0.1	0.2	0.3	0.4	0.5
Pedigree:					
Sire+dam	.22	.32	.39	.45	.50
Sire+dam + all grand parent	.27	.37	.43	.49	.53
All relatives	.29	.39	.45	.50	.54
Individual	.32	.45	.55	.63	.71
Full-sibs:					
5	.32	.42	.48	.53	.56
10	.41	.51	.56	.60	.62
20	.51	.58	.62	.65	.66
50	.60	.65	.67	.68	.69
100	.65	.68	.79	.69	.70
Half-sibs:					
5	.17	.23	.27	.30	.32
10	.23	.29	.33	.36	.38
20	.29	.36	.39	.41	.43
50	.37	.43	.45	.46	.47
100	.42	.44	.47	.48	.48
Progeny testing:					
5	.34	.46	.56	.60	.65
10	.45	.59	.70	.73	.77
20	.58	.72	.79	.83	.86
50	.75	.85	.90	.92	.94
100	.85	.92	.94	.96	.97

One record per animal, and environmental correlations are assumed to be zero.

11.11 Relative merits of the methods

The relative merits of some of the methods are illustrated in Figure 11.1 and summarised in table 11.4. Family and individual selection are of equal efficiency when the heritability is approximately 0.5 (Figure 11.1). When the heritability is lower family

selection is more efficient. When the heritability is higher than 0.5 individual selection is more efficient than family selection. The efficiency of family selection compared to individual selection increases markedly as the number per family increases whenever the heritability is below 0.4. Within family selection has low efficiency compared to all other selection methods according to Gall and Huang (1988a and 1988b).

The correlation between phenotypes within genetic groups is referred to as the intraclass correlation (t) and can be written as:

$$t = rh^2 + c^2 \qquad (11.2)$$

r= relationship between breeding values h^2 is heritability of the trait the and c^2 is the proportion of the total variance due to common environmental factors. If the heritability of the trait is unity as would be the case for a simply inherited trait like albino versus normal body colour and c^2 is zero then t = r. If h^2 is less than one, then $t = r\,h^2$, the case expected for quantitative traits when common environmental effects are absent.

12. DESIGN OF BREEDING PROGRAMS

BJARNE GJERDE

12.1 Introduction

The prospects of substantial genetic improvement have been well documented in several fish species (Gjedrem, 1997b). This potential should be exploited through effective and sustainable genetic improvement programs to develop strains for aquaculture with better performance, resource efficiency and product quality. Well-designed programs are also needed for the sustainable use of genetic resources as a safeguard against future changes in both production and market conditions (Hammond, 1994).

The term design may be used to characterise the size and structure of the breeding nucleus with respect to the number of full-sib families and breeding candidates tested per round of selection as well as the mating and selection strategies applied. An optimum designed program may be defined as one that maximise the genetic gain for a given trait (or the breeding goal) over a given period of time for a set of predefined constraints. Important constraints are the available testing capacity in terms of number of families (tanks) and breeding candidates that can be tested and the tolerable rate of inbreeding. Of importance could also be the establishing the base population, methods to trace pedigree (e.g., physical tags versus DNA tagging) and the degree of connectedness required across levels of fixed effects (e.g., cohorts, test environments, generations) to obtain unbiased estimates of breeding values and genetic gain. The design of the multiplier units is also important, but so far no study on this has been published.

The first large scale selective breeding program for farmed fish was set up for Atlantic salmon in the nineteen seventies (Gjedrem, 1992; Gjøen and Bentsen, 1997). The design of this program was based on basic knowledge in quantitative genetics, experiences from livestock programs and available technologies. Full-sib families were reared separately until tagging size and family identification was obtained by cold-branding and fin-clipping (Gunnes and Refstie, 1980). With the exception of some improvements (i.e. more traits in the breeding objective, more families tested and earlier and individual tagging using passive integrated transponder (PIT) tags) rather few changes have been taken place since then. The establishment of other family based programs for salmonids and other fish species, developed mainly in the late nineteen eighties and nineties, have to a large extent followed the same design.

The possibilities and constraints for the design of fish breeding programs have been discussed in general terms (Gall, 1990; Bentsen and Gjerde, 1994; Gjerde and Rye, 1997; Gjerde et al., 2002), but studies on optimum designs are few and limited to programs where truncation selection has been applied for a single normally distributed trait under individual (mass) (Gjerde et al., 1996; Villanueva et al., 1996; Bentsen and Olesen, 2002) and index selection (Villanueva and Woolliams, 1997; Trong, 2004).

T. Gjedrem (ed.), Selection and Breeding Programs in Aquaculture, 173–195.
© 2005 *Springer. Printed in the Netherlands.*

The design of breeding programs may be addressed from two different angles (Villanueva et al., 1996). *A posteriori*, the problem refers to methods of selection and evaluation for reducing inbreeding in the course of operation the breeding program. Such methods have received considerable attention in recent years (e.g. Toro and Perez-Enciso, 1990; Villanueva et al., 1994; Caballero et al., 1996; Meuwissen, 1997) and have been efficient in reducing rates of inbreeding with a minimal loss in genetic gain. However, they do not necessarily give maximum gains after a specified number of generations of selection. *A priori*, the problem is to maximum response given the basic design variables: available resources, traits and time scale objectives and attitude to risk (inbreeding). This will determine the value of additional resources, more of different measures or the expected benefits and losses from increased risks (Villanueva et al., 1996). Some of the proposed methods (e.g. different mating designs, non-random mating and optimum contribution selection) can be considered to be applicable to both problems.

The focus in this chapter will be on *a priori* methods and on the opportunities and challenges in designing fish breeding programs taking into account the reproductive characteristics of fish species (see Chapter 8) as well as recent and possible future technological developments that may influence the design of fish breeding programs.

12.2 Base population

The first step to start a breeding program is to collect the genetic material that forms the base population. In Norway, genetic material for Atlantic salmon was sampled from 40 river strains (Gjedrem et al., 1991) and no restriction was placed on strain contribution during the first generations. This resulted in an initial selection between strains. An alternative is to primarily mate animals from different strains before starting selection and apply low selection intensity during the first generations of selection. This strategy was used for the base population of a breeding program for Nile tilapia in the Philippines (Bentsen et al., 1997). This may secure the maintenance of a broad genetic variability (allelic diversity) that would allow long-term selection response and a stepwise inclusion of new traits in the breeding goal. In fish breeding, no study has focused so far on the effect of the design of the base population (i.e. the number of individuals to be sampled from one or several founder strains, their mixing, and the intensity of selection to be applied during the initial generations) on the magnitude and variability of the long-term selection response (risk) and inbreeding. It is worth noting that molecular measures of genetic diversity using neutral markers do not seem to be informative about the magnitude of quantitative genetic variation for traits of economic importance in a population (Reed and Frankham, 2001). Assessment of the magnitude of genetic variation for economic important traits applicable for farmed fish therefore should be obtained from traits recorded on fish raised in a commercial farm environment.

12.3 Maintenance of additive genetic variance

Breeding programs may be expected to be in place for many generations, especially for species with a short generation interval. Accumulation of inbreeding over generations reduces fitness in the population (e.g. reduced reproduction, viability and survival), reduces performance (inbreeding depression) of traits directly or indirectly (correlated responses) selected for and reduces the magnitude of the genetic variation thus jeopardising further genetic improvement. Consequently, maximization of short-term gains does not necessarily imply maximum long-term gains. Inbreeding also influences the variability of response, which determines the predictability of the future genetic mean of the population. The rate of inbreeding is thus an important parameter and has, as will be shown later, a major influence on the design of breeding programs.

The relation between the additive genetic variance in generation t+1 ($\sigma^2_{A_{t+1}}$) and the coefficient of inbreeding and the intensity and accuracy of selection may be written as (see Bulmer, 1985):

$$\sigma^2_{A_{t+1}} = \sigma^2_{s_{t+1}} + \sigma^2_{d_{t+1}} + \sigma^2_{w_{t+1}}$$

where

$\sigma^2_{s_{t+1}} = \sigma^2_{s_t}(1 - k_s r^2_{s_t}) = 0.25\sigma^2_{A_{t+1}}$ is the between sire genetic variance;

$\sigma^2_{d_{t+1}} = \sigma^2_{d_t}(1 - k_d r^2_{d_t}) = 0.25\sigma^2_{A_{t+1}}$ is the between dam (within sire) genetic variance;

$\sigma^2_{w_{t+1}} = \sigma^2_{w_0}(1 - F_t) = 0.50\sigma^2_{A_{t+1}}$ is the within full-sib family genetic variance;

where

$k_s = i_s(i_s - x)$ and $k_d = i_d(i_d - x)$ where i_s and i_d are the selection intensity for sires and dams, respectively and x is the standardized deviation of the truncation point from the population mean (Cochran, 1951); r_{s_t} and r_{d_t}, respectively, are the correlations between the true and predicted breeding values (accuracy of selection) for sires and dams in the previous generation; $\sigma^2_{w_0}$ is half of the additive genetic variance in the unselected base population and F_t is the inbreeding coefficient in the previous generation.

Thus, the between sire and the between dam (within sire) additive genetic variances are reduced each round of selection by a fraction $k_s r^2_{s_t}$ and $k_d r^2_{d_t}$, respectively (Pearson, 1903). This reduction, caused by the selection of parents, is due to negative covariance among genes at different loci in the same gamete; i.e. the joint gametic phase disequilibrium (Bulmer, 1971). In the current generation the within full-sib family additive genetic variance in the base population ($\sigma^2_{w_0}$) is reduced by a fraction equal to the inbreeding coefficient in the previous generation (Bulmer, 1971).

The disequilirium genetic variance (with unlinked loci) is reduced with one-half due to recombination of gametes in the progeny in each subsequent generation. This reduction,

however, does not go on indefinitely. A balance is soon reached where the reduction in the disequilirium variance due to selection is balanced by new variation created by recombination (Mendelian sampling variance). In the absence of linkage three or four generations of selection are sufficient to bring the population near to this balance. In the presence of linkage, the disequilibrium variance is reduced by a fraction less than one-half in each generation and thus it takes longer to reach the balance. The reduction in the genetic variance due to selection in the initial generations is often termed the Bulmer effect.

The proportion of the additive genetic variance in the base population remaining when the balance has been reached (the stabilized value of the utilizable additive genetic variance) decreases with increasing selection intensity and accuracy of selection. Fimland (1979) shows this for different values of the intensity and accuracies of selection when an infinitesimal genetic model and infinite number of individuals in the breeding population is assumed (i.e., no accumulation of inbreeding). Assuming a finite population and a rate of inbreeding of ΔF=0.5 % or ΔF=1% per generation, Figure 12.1 shows the utilizable additive genetic variance over 10 generation of selection for four combinations of the intensity and accuracies of selection. These results illustrate the importance of maintaining the additive genetic variance in the base population by applying moderate selection intensities and rates of inbreeding. Of importance is that a 50 % increase in the accuracy of selection results in a much larger reduction in the genetic variance than a 50 % increase in the selection intensity.

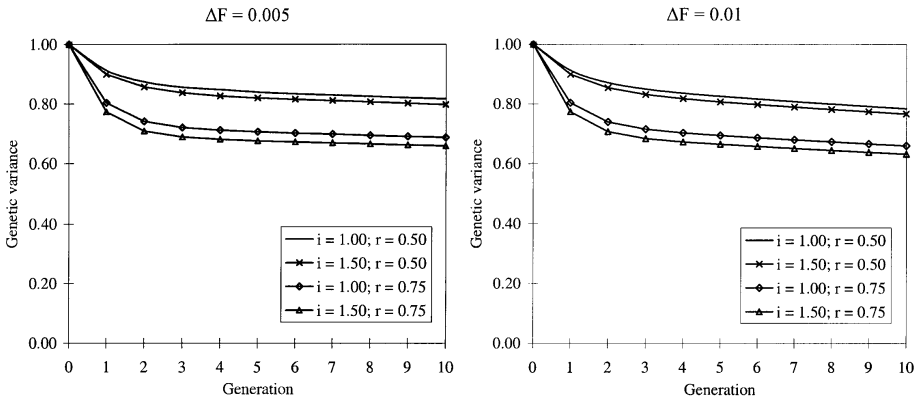


Figure 12.1 Reduction in additive genetic variance over 10 generations of selection for four combinations of the intensities and accuracy of selection at two rates of inbreeding (ΔF=0.005 and 0.01).

12.4 Inbreeding and risk

When no selection, no mutation and random mating is assumed, the rate of inbreeding per generation in a closed population can be calculated as $\Delta F = 1/2N_e = 1/8N_s + 1/8N_d$; where N_e is the effective population size and N_s and N_d, respectively, are the number of

male and female parents (Wright, 1931). When selection is practised the above formula may grossly underestimate the rate of inbreeding (see below).

Meuwissen and Woolliams (1994a) investigated the effective population sizes required for livestock populations to avoid a decrease in fitness. Assuming a zero correlation between fitness and a production trait under artificial selection, they found that the minimum effective population size required to avoid a decline in fitness after ten generations of selection ranges from N_e=31 (ΔF=1.6%) to N_e=250 (ΔF=0.2%) per generation, depending on the magnitude of the additive genetic variance of fitness ($N_e = D/2\sigma^2_{wa}$, where D=the inbreeding depression of fitness with complete inbreeding, and σ^2_{wa} = the additive genetic variance of fitness). Thus higher effective population sizes are required for low than for high values of the heritability of fitness. If fitness declines as a correlated negative response to artificial selection, then a large increase in the critical effective population size is required (Meuwissen and Woolliams, 1994a).

Distinctive reproductive characteristics in fish allow for accurate estimation of breeding values and for high selection intensities and therefore high short-term selection responses. However, high selection intensities and accuracies may result in rapid accumulation of inbreeding and this can jeopardise further genetic improvement, reduce fitness and performance and also produce highly variable responses (Meuwissen and Woolliams, 1994b). Implementation of measures for restricting inbreeding is therefore essential in fish breeding programs.

Strategies to control the rate of inbreeding may include increasing the size of the breeding population, increasing the number of parents selected, limiting the use of each parent, limiting the number of individuals selected from each family and choice of mating design (e.g. factorial versus nested design; Woolliams, 1989). However, in a practical situation such procedures may have to take place for a given set of facilities. For instance, in a situation where full-sib families are reared separately until a size suitable of tagging and for a given mating system, the number of parents selected cannot be increased unless more tanks are made available.

A series of methodologies have been developed for finding optimal designs of breeding programs that give maximum genetic gains with constraints on inbreeding in schemes with discrete generations under individual selection (Gjerde et al., 1996; Villanueva et al., 1996, 2000), index (Villanueva and Woolliams, 1997) and BLUP (Meuwissen, 1997; Grundy et al., 1998) or in schemes with overlapping generations under BLUP selection (Meuwissen and Sonesson, 1998; Grundy et al., 2000). Some of these methods (Meuwissen, 1997; Grundy et al., 1998; Meuwissen and Sonesson, 1998; Grundy et al., 2000) are dynamic, in that they adapt to current selection candidates (i.e., optimum contribution selection), and can therefore correct skewness in contribution of families over generations. Skewness of contributions by families results in higher rates of inbreeding.

The optimisation tools generally deal with normally distributed traits that can be

measured on live breeding candidates. Many traits of economic importance in fish do not fulfil these requirements (e.g. survival, carcass quality and disease resistance traits). Thus there is need to investigate and possibly extend the current theory to provide selection tools for scenarios relevant to fish species.

Most current fish breeding programs have closed breeding nuclei. Low levels of migration (i.e., an open breeding nucleus strategy) are effective to break up inbreeding and reintroduce genetic variation (Falconer and MacKay, 1996). The migrants could come from selected populations with similar genetic performance. However, introduction from inferior (wild) populations is often the only option for the introduction of new genetic variation (Osman and Robertson, 1968). In such instances it is difficult to maintain the improved performance of the improved population. New genetic variants also arise infrequently by mutation. The balance between new genetic variation by mutation and loss of genetic variation due to inbreeding will depend on the relative rates of mutation and inbreeding (see Chapter 2).

Genetic gain is often achieved at increased risk (variance of response). There is a strong relationship between inbreeding and variance of response, such that addressing one of these two components of risk usually suffices and results indicate that large reductions of variance of response and rates of inbreeding can be combined with a small reduction in genetic gain (Meuwissen and Woolliams, 1994a). Nicholas (1989) suggested that to be confident that the expected response will actually be achieved in practice, acceptable coefficients of variation of genetic gain per generation should be not higher than 10% after ten years of selection.

12.5 Mating design

A prerequisite for running a sustainable genetic improvement program is that a reasonably high number of full- and half-sib families can be produced in a controlled and reliable manner. The possibility for separate collection of eggs and semen in many fish species facilities a wide range of mating designs. In some species (e.g. salmonids) a high number of large full-sib and maternal and/or paternal half-sib groups may be produced from a set of simultaneously artificially stripped breeders and spawning may often be synchronised or induced.

For some species artificial stripping is difficult to perform or it may be difficult to obtain a sufficient number of synchronized breeders that can be stripped and mated. For these species the production of full- and half-sib families may be obtained through natural spawning by keeping pairs of fish in separate tanks. This has proven successful in Atlantic cod (Terjesen et al., 2004). The newly fertilized eggs are collected through the outlet of each tank. Paternal or maternal half-sib families were obtained, by first mating one pair of parents, then replacing one of the parents in the tank.

Choice of the most appropriate mating design is dependent on several factors. Most important is whether stripping or natural mating can be performed, the availability of sexual mature breeders of each sex and also the type of genetic effects (i.e., additive and

non-additive genetic) for which unbiased and accurate parameter estimates or predictions of breeding values are to be obtained. For example, significant non-additive genetic variation for harvest body weight has been detected within fish populations (Rye and Mao, 1998; Pante et al., 2002) and in such a case factorial or mixed nested/factorial may provide more accurate estimates of both additive and non-additive genetic effects than can be obtainable in a purely nested design.

12.5.1 MASS SPAWNING

Natural mass spawning in which several sexual mature male and female breeders are kept in the same tank is often used when producing grow-out animals, in particular for species where artificial stripping is difficult to perform. However, the relative contribution of each breeder to the total number of offspring in the breeding population will be unknown. If individual selection is to be applied, the selected breeders may be the offspring from a rather limited number of parents. Therefore, the effective population size may become low and consequently the rate of inbreeding high as demonstrated in the gilthead sea bream (*Sparus aurata*) where the effective population size, calculated from the number of offspring sampled from a single day spawning, was less than one third of the actual number of parents used in the mass spawning (Brown et al., 2004). However, the use of DNA markers for parental assignment may overcome this problem (see section 12.9). Use of DNA markers for parental assignment in mass spawning design may also facilitate the use of sib records in the selection decisions. However, due to the unknown and most likely very variable number of offspring from each parent, the number of breeding candidates needed to be genotyped may become high (section 12.9). Consequently, natural mass spawning is not recommended in the breeding nucleus.

12.5.2 SINGLE PAIRS

This is the simplest mating design in which the milt from each male is used to fertilise the eggs from only one female (Table 12.1). Only full-sibs are produced, the number being equal to the number of males or females. In this design additive genetic and other effects common to full-sibs (non-additive genetic effects and maternal and common environmental effects) are confounded and cannot be assessed independently. This may result in biased breeding values with low accuracy, resulting in low genetic gain. Consequently, single pair mating should not be used unless the non-additive genetic effects and the effects common to full-sibs other than additive genetics are low.

Table 12.1. Single paired mating design, 1♂ : 1♀

	Female		
Male	1	2	3
1	x		
2		x	
3			x

Full-sib model: $y_{fik} = fixed_f + fs_i + e_{fik}$

$$\sigma^2_{fs} = 0.5\sigma^2_A + 0.25\sigma^2_D + \sigma^2_M + \sigma^2_C$$

Animal model (with A-matrix): $y_{fijk} = fixed_f + a_i + fs_j + e_{fijk}$

$$\sigma^2_a = \sigma^2_A \; ; \; \sigma^2_{fs} = 0.25\sigma^2_D + \sigma^2_M + \sigma^2_C$$

12.5.3 NESTED

The most commonly used mating design in today's fish breeding programs is the nested design. Milt from one male is used to fertilise eggs from several females (females nested within males, Table 12.2a). Thus both full- and paternal half-sibs are produced. The number of full-sib groups is equal to the numbers of females and the number of half-sib groups is equal to the number of males. In this design the sire and dam component of variance each account for one quarter of the additive genetic variance. However, the dam component may be inflated by non-additive genetic (dominance), maternal (e.g. caused by different egg sizes among the female breeders and possible cytoplasmatic genetic effects) and common environmental effects caused by separate rearing of full-sib families until to tagging.

Table 12.2a Mating design where females are nested within male; 1♂ : 2♀ and 1♂ : 3♀

	Female				
Male	1	2	3	4	5
1	x	x			
2			x	x	x

Sire and Dam model: $y_{fijk} = fixed_f + s_i + d_{ij} + e_{fijk}$

$$\sigma^2_s = 0.25\sigma^2_A \; ; \; \sigma^2_d = 0.25\sigma^2_A + 0.25\sigma^2_D + \sigma^2_M + \sigma^2_C$$

Table 12.2b Mating design where males are nested within female; 1♀ : 2♂ and 1♀ : 3♂

	Male				
Female	1	2	3	4	5
1	x	x			
2			x	x	x

Dam and Sire model: $y_{fijk} = fixed_f + d_i + s_{ij} + e_{fijk}$

$$\sigma^2_d = 0.25\sigma^2_A + \sigma^2_M \; ; \; \sigma^2_s = 0.25\sigma^2_A + 0.25\sigma^2_D + \sigma^2_C$$

Alternatively, the eggs from one female may be divided into portions and each portion fertilised with milt from different males (males nested within females, Table 12.2b). Thus both full- and maternal half-sibs are produced. This design may be a good alternative, particularly in cases where the availability of an adequate number of female breeders is less than that of male breeders. The number of full-sib families is equal to the number of males and the number of half-sib families equal to the number of females. In this design the dam component of variance accounts for a quarter of the

additive genetic variance but may be inflated by maternal effects. Consequently, this design will in general provide a less accurate estimate of the additive genetic variance in the population than the sire component estimate derived from the design in Table 12.2a.

12.5.4 FULL FACTORIAL

Semen from each male is used to fertilise eggs from several females and, at the same time eggs from each female are divided into portions and fertilised with semen from different males (Table 12.3). Thus both full-sibs as well as paternal and maternal half-sibs are produced. For each set produced the number of full-sib groups is equal to the number of males multiplied with the number of females. The number of paternal half-sib groups is equal to the number of males and the number of maternal half-sib groups is equal to the number of females. This is a good design to obtain reliable estimates of both additive and non-additive genetic variance in the population. However, a serious drawback is that a low number of male and females will be tested for a given number of rearing tanks available (see Table 12.5).

Table 12.3 Factorial mating design; 2 : 2 and 3 : 3

	Female				
Male	1	2	3	4	5
1	x	x			
2	x	x			
3			x	x	x
4			x	x	x
5			x	x	x

Animal model (with A-matrix): $y_{fijk} = fixed_f + a_i + fs_j + e_{fijk}$

$\sigma^2_a = \sigma^2_A$; $\sigma^2_{fs} = 0.25\sigma^2_D + \sigma^2_M + \sigma^2_C$

Animal model (with A- and D-matrix): $y_{fijk} = fixed_f + a_i + d_j + fs_j + e_{fijk}$

$\sigma^2_a = \sigma^2_A$; $\sigma^2_d = 0.25\sigma^2_D$; $\sigma^2_{fs} = \sigma^2_M + \sigma^2_C$

12.5.5 PARTLY FACTORIAL

This design (Table 12.4) may be a good alternative to a nested or factorial design. For a given number of available tanks, a higher number of male and female breeders will be tested than in a factorial design while at the same time obtaining better connectedness between the full- and half-sib families than in a nested design. In the presence of a tank effect, genetic parameter estimates (i.e., heritability) with higher accuracy and precision were obtained from a partly factorial than from a nested design (Berg and Henryon, 1998). Most likely this will also be the case for the estimation of non-additive genetic effects.

Table 12.4 Partly factorial mating design

Male	Female				
	1	2	3	4	5
1	x	x			
2		x	x		
3			x	x	
4				x	x
5	x				x

Same models as for Table 12.3

12.5.6 FUTURE MATING DESIGN

Use of optimum contribution selection (e.g., Meuwissen, 1997) implies that not a specific mating design will be used. Rather the output from this selection algorithm will give information about how many offspring (or mates when assuming an equal number of offspring per mating) a given animal should get to maximize the genetic gain while at the same keep the rate of inbreeding at an accepted level. Thus the developments of new selection tools may change the above strict systematic mating design to more mixed designs (see also section 12.7).

12.5.7 MATING DESIGN AND INBREEDING

For a given testing capacity (i.e., number of full-sib families that can be reared separately until tagging) different mating designs allow for different number of sires and dams to be tested (Table 12.5). If a given number of scored selection candidates and truncation selection is assumed, the different mating design will result in different effective population sizes and rates of inbreeding. Assuming no selection and random mating, this is illustrated in Table 12.5. The lowest rate of inbreeding is obtained with the paired mating design. For nested mating designs the rate of inbreeding increases with increasingly skewed sex ratio. The mixed nested/factorial and the pure 2 x 2 factorial designs give the same rate of inbreeding. Factorial designs with each sire (dam) mated to more than two dams (sires) result in fewer sires and dams tested and thus in higher rates of inbreeding. When selection was applied in MOET (Multiple Ovulation and Embryo Transfer) dairy cattle schemes, factorial design resulted in lower rates of inbreeding than nested design without reduction in genetic gain (Woolliams, 1989; Villanueva et al., 1994). Studies on appropriate mating design in fish breeding programs are required.

12.6 Selection

12.6.1 INDIVIDUAL SELECTION

Gjerde et al. (1996) investigated optimum designs for individual selection programs in fish breeding under constrained inbreeding through stochastic simulation. For a given number of scored selection candidates (N = 300 to 9600 animals), and mating ratio (d =

2, 6 and 10 dams per sire), rates of genetic gain and inbreeding were obtained separately for different number of sires. Designs given a rate of inbreeding close to the specific value (ΔF = 2%; 1%, 0.5% and 0.25%; corresponding to N_e = 25, 50, 100 and 200, respectively) were defined as optimum.

Table 12.5 Number of sires and dams, effective population size (N_e) and rate of inbreeding per generation (ΔF). Assumed are no selection, random mating and a testing capacity of 100 full-sib family groups per generation.

Mating design[1]	Sires or Dams	Dams or Sire	N_e[2]	ΔF, %[3]
Paired				
1 : 1	100	100	200	0.25
Nested				
1 : 2	50	100	133	0.38
1 : 3	33	100	100	0.50
1 : 4	25	100	80	0.63
Factorial				
2 x 2	50	50	100	0.50
3 x 3	33	33	66	0.75
4 x 4	25	25	50	1.00
Nested / Factorial				
2 : 2	50	50	100	0.50

[1] Male to females, or female to males, ratio

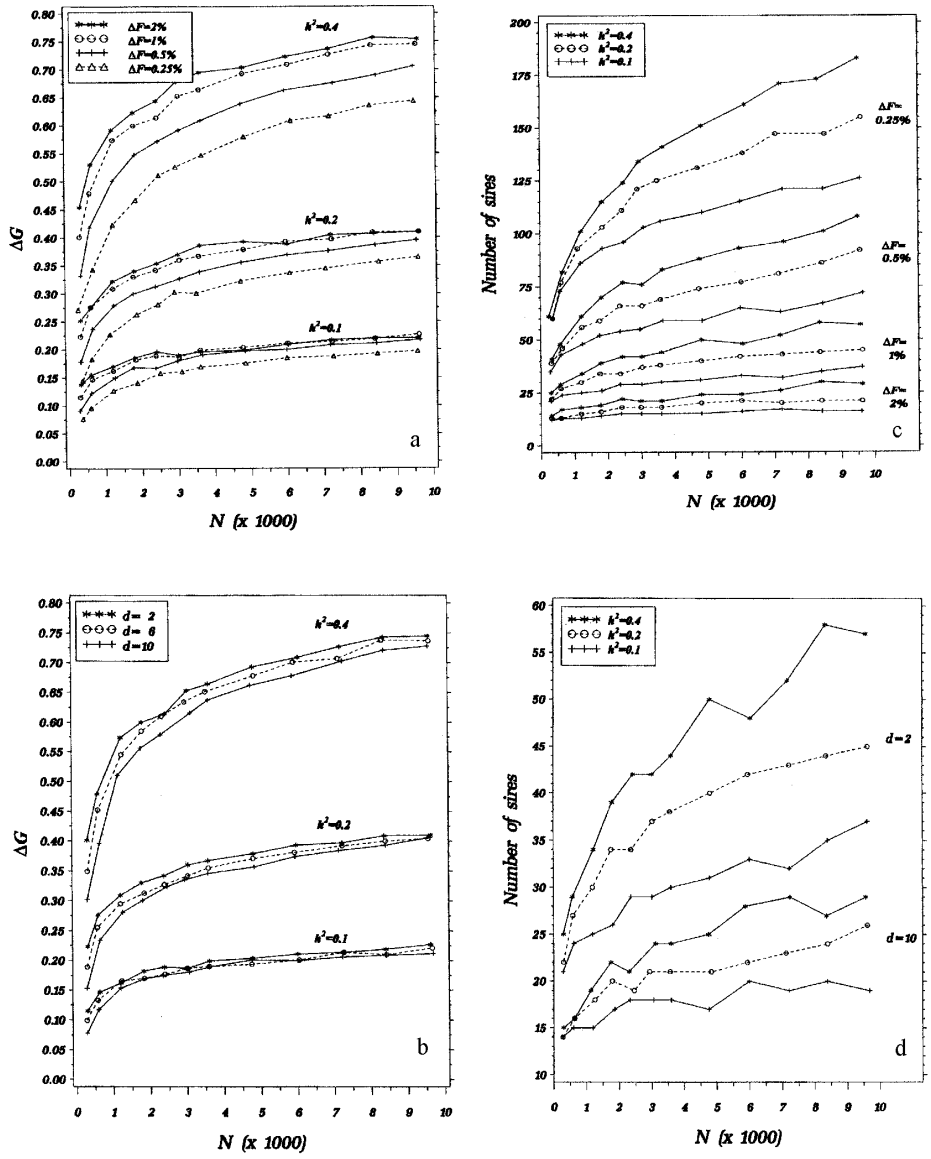
[2] $N_e = \frac{1}{2\Delta F} = \frac{4N_m N_f}{N_m + N_f}$; [3] $\Delta F = \frac{1}{2N_e}$

As expected the genetic gain increases with increasing population sizes, in particular for small populations and high heritabilities (Figure 12.2a). Imposing a constraint on inbreeding resulted in lower genetic gain, in particular for high heritabilities (Figure 12.2a).

The desired rate of inbreeding has a large influence on the design, in particular for high heritabilities (Figure 12.2b). Thus for the largest population size studied (N = 9600) and for h^2 = 0.2 and d = 2, the optimum design uses approximately 35 sires (and 70 dams and 274 animal per full-sib family) for ΔF = 1% while 90 sires were used (and 180 dams and 53 animal per full-sib family) for ΔF = 0.5%.

The mating ratio has only a minor effect on the genetic gain (Figure 12.2c) but a large influence on the design, in particular for high heritabilities (Figure 12.2d). Thus for N = 9600, h^2=0.2 and ΔF = 1% the optimum design was approximately 45 sires (and 90 dams and 107 animal per full-sib family) for d=2 while it was 25 sires (and 250 dams

and 38 animal per full-sib family) for d = 10. The effect of lower inbreeding rates and



Figrue 12.2 Predicted genetic gain (ΔG) per generation (Figure 12.2a and 12.2c, d=2), and optimum number of sires (Figure 12.2b and 12.2d) for ΔF=1%, when applying individual selection for different population sizes (N), heritabilities (h^2), rates of inbreeding (ΔF) and mating ratios (d=number of dams mated to each sire).

higher mating ratios is that more but smaller full-sib families must be produced. In the

above-mentioned study (Gjerde et al., 1996), for a given set of value of N, ΔF, d and h^2, several simulations must be run to find the optimum solution. The number of selected animals is changed in different simulations until the desired level of inbreeding is achieved. This procedure is very computationally demanding, in particular for large population sizes.

12.6.2 WITHIN FAMILY SELECTION

When selection is practised only for trait(s) that can be recorded on the live breeding candidates, a within-family selection protocol (called 'walk-back' selection, Doyle and Herbinger, 1994) has been proposed. In this protocol, DNA markers are used to identify the pedigree of the potential breeders: First all breeding candidates are graded with respect to the trait(s) selected for. Then the most superior (e.g. the largest) animal is DNA fingerprinted and chosen to become a parent of the next generation. Then the second-largest animal is genotyped and added if it comes from a different family than the first, otherwise it is discarded. The third-largest individual is fingerprinted and accepted if it is not a full-sib of the first two. This process is repeated until a sufficient number of pairs of selected breeders, each from a different full-sib family, are obtained. The protocol maximizes the effective population size, but is most likely not to be optimal in terms of genetic gain. Also, with large and variable family sizes the number of fish that need to be genotyped may become very high with a correspondingly high cost. Because the probability of selecting related animals increases with increasing heritability, the number of animals to be genotyped will be higher when selection is practised for trait(s) with high than for trait(s) with low heritability.

Recently the 'walk-back' procedure has been further developed to be applicable also for individual selection (Trong, 2004). If the average coancestry (additive genetic relationship) among the selected candidates (i.e., the inbreeding level in the next generation) is found to be too high, additional fish are selected and their genetic relationship identified. This procedure is repeated until the desired level of relatedness among the breeders is accomplished. This may be further developed to also utilize sib information.

12.6.3 INDIVIDUAL VERSUS BLUP SELECTION

Gjøen and Gjerde (1998) used stochastic simulation to find optimum design of breeding programs when applying truncation selection on BLUP (BTS) or phenotypic values (PTS). When the constraint on inbreeding was set equal to $\Delta F = 1$ % per generation, genetic gain from BTS and PTS was similar for $h^2 = 0.10$, but PTS produced higher genetic gain than BTS for $h^2 = 0.20$ and $h^2 = 0.40$ (Figure 12.3). The reason is the higher selection intensity for PTS more than compensates for the higher accuracy of selection for BTS (Table 12.6). The breeding design is quite different for the two selection methods with BTS having substantially large number of full- and half sib families, but the size of each of these families is smaller (Table 12.6). Consequently the costs of BTS and PTS schemes may be quite different, in particularly if the full-sib families of the BTS scheme have to be reared separately until a size suitable for tagging.

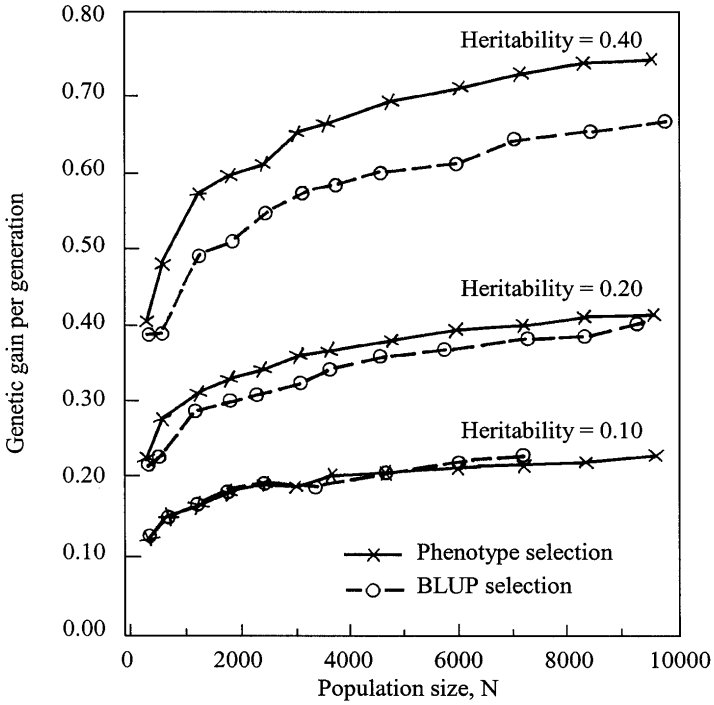


Figure 12.3 Predicted genetic gain for different population sizes (N) and heritabilities (h^2) for phenotype (individual) and BLUP truncation selection ($\Delta F = 1\%$, d = 2 dams per sire). Reproduced from Gjøen and Gjerde (1998) by permission of Organising Committee.

A special BTS scheme, where costs of rearing the fish until tagging will be the same as for PTS schemes, occurs when candidates are pooled at fertilization or shortly thereafter and DNA markers are used to assign offspring to parents (Sonesson et al., 2004). After tagging, the cost of rearing fish should be the same for BTS and PTS provided the same number of fish are tagged.

Table 12.6. Optimum mating design and genetic gain when applying truncation selection on phenotypic (PTS) or BLUP (BTS) values; N = 600 males + 600 females, $h^2 = 0.40$, mating ratio 1♂ : 2♀ and ΔF=1%. Modified from Gjøen and Gjerde (1998) by permission of Organising Committee.

	PTS[1]	BTS[2]
No. of sires	34	103
No. of dams	68	206
No. offspring per full-sib family	18	6
Intensity of selection (i)	1.90	1.15
Accuracy of selection (r)	0.56	0.76
Additive genetic variance	0.54	0.56
Genetic gain per generation	0.57	0.49

[1] Own phenotype information; [2] Own, full- and half-sib information

Thus it may be argued that comparison of schemes should be done with a constraint on the number of tanks to rear the full-sib families as was done by Sonesson et al. (2004). These authors found that for the same number of selected sires and dams, the rate of inbreeding was much higher for the BTS than for the PTS schemes, while genetic gain was quite similar (Figure 12.4). At the same level of inbreeding the design of the BTS and PTS scheme was quite different with BTS having substantially large number of full- and half sib families, but the size of each of these families is smaller. These results are in agreement with those presented in Table 12.5.

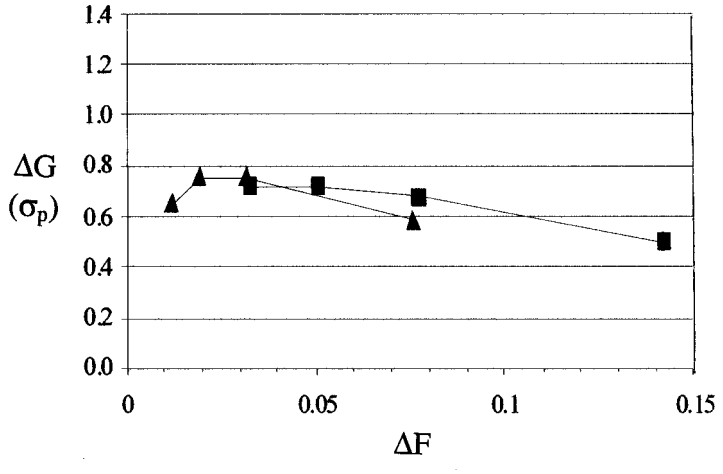


Figure 12.4. Rate of inbreeding (ΔF) and genetic gain (ΔG) for PTS (triangles) and BTS (squares) selection for a trait with h^2=0.4. Schemes with 100 tanks, 6400 selection candidates and 100, 50, 20 and 5 selected sires and dams going from left to right in the graph were used. Reproduced from Sonesson et al. (2004) by permission of AKVAFORSK.

As an aside, the danger of using the simple formula of Wright (1931) can be demonstrated by the results in Table 12.5 where the rate of inbreeding, as measured by the pedigree, was restricted to be 1 % per generation. If the simple formula of Wright (1931) is used to calculates the rate of inbreeding, the predicted rate of inbreeding was 0.55 % for PTS and 0.18 % for BTS; thus a substantial underestimating of the rate of inbreeding and more so for BTS for PTS.

12.6.4 SIB SELECTION

When traits cannot be recorded on the live breeding candidates (e.g., carcass quality and disease resistance traits), selection within families cannot be applied and consequently the selection intensity obtained is that between full- and half-sib families (sib selection). The lack of appropriate technologies for recording such traits represent a challenge in fish breeding programs and may have a substantial effect on the design of the programs. When such technologies become available, the same genetic gain could be obtained with a lower number of selection candidates and families, or higher gains could be

obtained with the same number of individuals and families. Studies on the optimum design of fish breeding programs when applying sib selection has not been undertaken. The development of technologies to record such traits should be encouraged.

12.6.5 PROGENY TESTING

Most fish species, as opposed to salmonids, are multiple spawners making progeny testing of both males and female breeders possible. For male breeders of any species the same could be obtained if reliable procedure for cryopreservation of milt is available. Selection could thus be performed among the parents based on their progeny as well as own and sib records. The highest ranking could be used to produce a second crop of progeny, either for use in the nucleus or for the production of grow-out animals. Optimum mating design utilising progeny test records will require that each male (female) be mated to more females (males) than when utilising only own and/or sib records (see Figure 13.2).

12.7 Selection and mating

Optimisation tools have mainly dealt with selection decisions, assuming that the selected parents are mated at random and using a given mating design. The reproductive characteristics of fish imply a high potential for non-random mating strategies that have not been investigated to date and few attempts have been made to optimise simultaneously selection and mating decisions. However, a number of non-random mating systems have been proposed to control inbreeding, implying that their application could lead to higher gains without increases in inbreeding. This may include compensatory mating (Caballero et al., 1996), minimum (Toro et al., 1988) and average co-ancestry mating applicable for discrete (Meuwissen, 1997; Kerr et al., 1998; Sonesson and Meuwissen, 2000) as well as overlapping generations (Meuwissen and Sonesson, 1998).

Use of optimum contribution selection (e.g., Meuwissen, 1997; Grundy et al., 1998; Meuwissen and Sonesson, 1998; Grundy et al., 2000) implies that not a specific mating design will be used. Rather the output from this selection algorithm will give information about how many offspring (or mates when assuming an equal number of offspring per mating) a given animal should get to maximize the genetic gain while at the same keep the rate of inbreeding at an accepted level. For low number of selection candidates (i.e., ≤ 512 newborn candidates per year) optimum contribution selection has given equal or higher genetc gain than BLUP selection and with a substantially lower number of parents selected (Meuwissen, 1997; Meuwissen and Sonesson, 1998). In traditional farm animals the optimum number of offspring may be difficult to obtain because of reproductive limitations, particularly in females. However, in fish species such reproductive limitations are in general less, and therefore optimum contribution selection is of particular interest for fish species. The application of this new selection tool for fish species should be studied and in particular for larger population sizes than done so far. Its influence on the accuracy and unbiasedness of genetic parameters should also be investigated.

12.8 Connectedness

Breeding programs need efficient procedures to obtain accurate and unbiased estimates of breeding values to maximise genetic gain and to obtain unbiased estimates of genetic changes. Unbiased estimates can only be obtained if environmental and genetic differences across levels of fixed effects are accounted for (Sorensen and Kennedy, 1984a; 1984b). BLUP methodology offers this possibility provided that there is sufficient connectedness in the data (Kennedy and Trus, 1993; Hanocq et al., 1996). Studies on the level of connectedness needed in fish programs for unbiased estimates of breeding values have not yet been undertaken.

In fish breeding programs the question of appropriate level of connectedness arises when (a) not all families are reared at all test units in order to save costs; (b) the breeding population consists of several cohorts of families produced at different times of the year; (c) genetic material is exchanged between different breeding nuclei to reduce inbreeding and to increase the intensity of selection; and (d) sires and/or dams are reused over generations (e.g. for monitoring genetic changes).

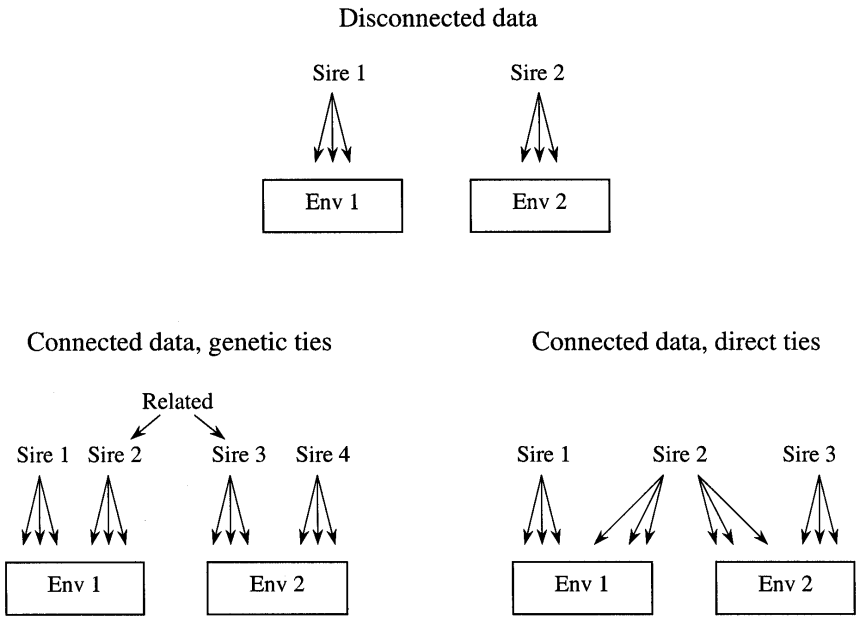


Figure 12.5 Illustration of data that are disconnected and data that are connected through genetic ties or direct ties.

Connected data can be obtained either from parent-offspring relationships (genetic ties) or from parents having offspring with records (direct ties) across different levels of fixed effects (e.g. within and across generations). An illustration of this is given in Figure 12.5.

In some species it may be possible to create an adequate number of direct ties by keeping randomly bred control groups, or by repeated mating of parental pairs that have been used in the past. Maintaining stable control groups requires substantial resources in terms of testing capacity (Gow and Fairfull, 1990 and Chapter 15). Repeated mating requires that both sires and dams are kept alive over generations, or that cryopreserved embryos will be available. This strategy may also give biased estimates of gain unless possible non-additive genetic effects can be accounted for. The alternative is to reuse some sires (here cryopreserved sperm could be used) or dams only.

12.9 Physical tags versus DNA-markers for parental assignment

By using DNA (genetic) markers for parental assignment in fish breeding programs, the pooling of families at fertilisation or shortly thereafter would be possible. This could eliminate the need for costly multitank facilities and the problem of confounding genetic and common environmental effects (Doyle and Herbinger, 1994). However, the use of genetic markers has some potential dangers primarily as a result of low and variable survival rates among the families from the time of pooling the families to the time of selection of parents. Consequently, the number of fish that need to be genotyped may become very high in order to keep the rate of inbreeding at an acceptable level in programs that apply individual selection, and to obtain records on a sufficient number of animals from all families in programs that utilise sib information in the selection decisions.

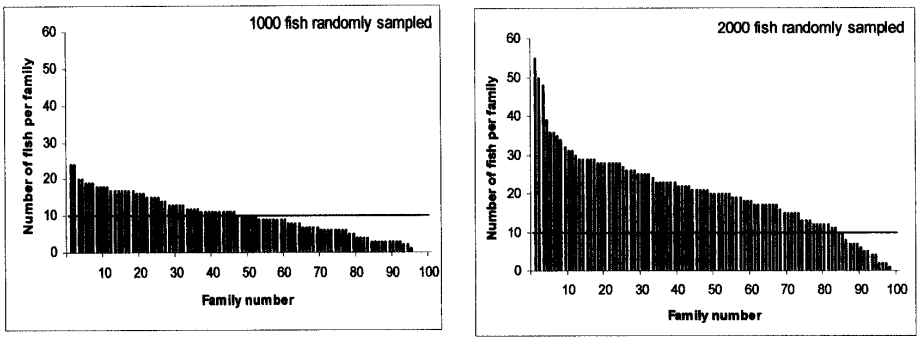


Figure 12.6 Number of fish genotyped per full-sib family when sampling randomly 1000 and 2000 fish from a population of 10000 surviving fish (500 fish pooled from each of 100 full-sib families; mean survival rates 20%, SD =10% of the family mean survival rates).

This is illustrated in Figure 12.6 where only 52 and 85 out of 100 families had 10 or more fish per family when genotyping a random sample of 1000 and 2000 individuals, respectively. Some families had no individuals genotyped while other families were represented with a much higher numbers of individual than required to obtain a breeding value with high accuracy. Thus the nucleus breeding population can become

unbalanced leading to decreased selection intensities and accuracies and thus lower genetic gain.

The number of DNA markers characterised for commercial species are sufficient for correctly assigning offspring to parents (e.g., Norris et al., 2000; Villanueva et al., 2002) but the technology is still very costly, particularly given the fact that fish need to be retyped each time family assignment is required (e.g. if trait records are obtained at different times). The combined use of DNA markers and physical tags can overcome this problem to some extent. Also the breeding candidates need to be physically tagged so that those to be selected can be easily traced at time of mating.

Nevertheless, there are a few companies that use genetic markers in their breeding programs. Further improvement in genotyping technology may reduce the costs considerably. Therefore it is of interest to evaluate the use of genetic markers for parental assignment in fish breeding programs. Using DNA markers may not only reduce the costs of the programs but also increase the genetic gain.

In addition, to remove the need for separate rearing, use of DNA markers provides the opportunity to (1) increase the numbers of individuals and families without additional investment in units for separate rearing (increasing thus the scope for genetic improvement), (2) to compute exact genetic relationships among siblings (Norris et al., 2000; McDonald et al., 2004) and (3) to control rate of inbreeding in programs applying individual selection (Doyle and Herbinger, 1994; Trong, 2004). It would also be possible to include markers linked to loci affecting economically important traits as part of the panel of markers used for parental assignment.

12.10 Small scale breeding program

For less important economic species, simple programs that utilise only individual selection primarily for growth should be initiated. These programs may later be extended to more advanced programs that also utilize sib information in the selection decisions if production and thus the economic importance of the species increases. Since individuals are not tagged when applying individual selection, the choice of mating and breeding design structures that gives tolerable rates of inbreeding is of particular importance for such programs. Key issues for such programs are the use of a sufficient number of breeders and restricting the number of sibs tested per family (Gjerde et al., 1996, Bentsen and Olesen, 2002; Figure 12.2).

The inbreeding level in the grow-out animals may be kept very low if breeders from different populations or lines are mated to produce the grow-out animals. In the grow-out phase the individuals from the lines may be kept in the same rearing unit but need to be identified, for example by fin-clipping. Thus an additional line will not require much extra investment and resources.

To account for possible genotype by environment interaction, a sample of the breeding candidates may be raised at different locations representing different commercial farm

environments. Some breeders must then be selected from each of these test environments (see Chapter 16).

12.11 Large scale breeding program

Breeding programs for important economic species should utilize both individual and/or family information in the selection of breeders. The design and logistics of a typical large-scale family based fish breeding program are shown in Figure 12.7.

12.11.1 BREEDING NUCLEUS

The breeding nucleus typically consists of a large number of full-sib and paternal or maternal half-sib family groups, depending on the mating structure. Generations are usually discrete but in some species the use of light regimes makes it possible to produce cohorts of families at regular intervals throughout the year, leading to overlapping generations. Currently, the full-sib families are kept in separate units/tanks until a size suitable for tagging by physical methods. A sample of individuals from each full-sib family are tagged and raised as breeding candidates preferably in a commercial farm environment.

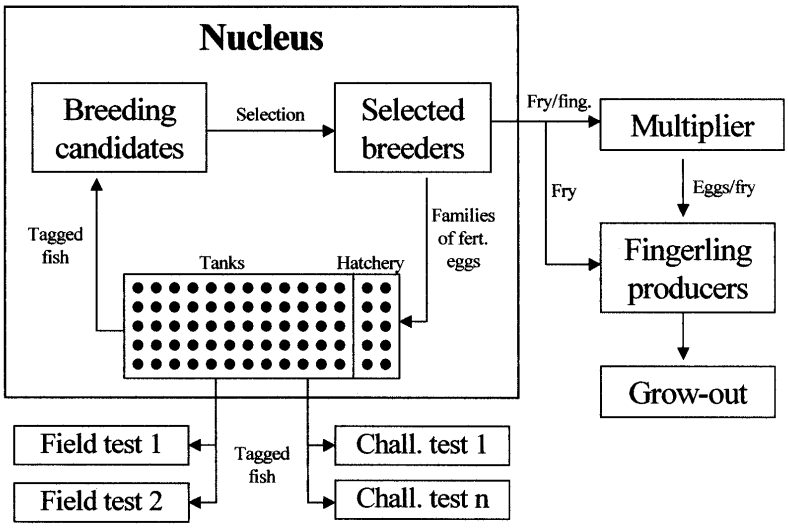


Figure 12.7 Organogram showing the main elements in a fish breeding program.

The high fecundity of fish allows the use of data from full-sib families that are replicated in facilities outside the breeding nucleus, in the genetic evaluations. The nucleus operations could then focus solely on the production, tagging and distribution of test fingerlings from a large number of families. This would help protect the nucleus

against accidents, disease breakdowns and would secure the necessary number of individuals to be able to carry out effective selection (Bentsen and Gjerde, 1994). In this case, the breeding candidates could be raised in any type of environment producing quality gametes. Such a strategy will imply sib selection for all of the traits that are selected, including those that can in theory be recorded on live breeding candidates. But most likely growth would still be recorded on the candidates.

12.11.2 TEST ANIMALS

Other samples of tagged individuals from each family are raised as test animals in commercial farm environments either at facilities where destructive carcass quality traits can be recorded or at facilities at which disease challenge tests can be performed. Performance data from different commercial farm environments can be useful for selecting fish that are robust to differences in environmental conditions, provided the magnitude of genotype by environment interaction is marginal and does not lead to severe reranking of families in different environments. Sufficiently connected data (see above) could be obtained with only some of the families reared in each test environment, saving thus some costs.

Studies are needed to determine the minimum number of farm environments needed for eliminating the risk of producing animals only suited for a particular type of farm environment. This will of course depend on the magnitude of the genotype by environment interaction.

12.11.3 SELECTION

Based on the data recorded on the test animals and the breeding candidates, animals with high breeding values are selected among the breeding candidates to become the breeders for the next generation. After mating the resulting full-sib families are transferred to separate units in the hatchery and thereafter to separate tanks for growing until tagging size.

12.11.4 MULTIPLIER UNITS

The improved genetic material (surplus milt and eggs from the selected breeders or from other animals with high breeding values) are transferred as fertilized eggs to one or several multiplier units where the animals are raised to sexual maturity to produce eggs/fry that are sold to fingerling producers. Alternatively, fry/fingerlings are produced and are directly supplied from the selected nucleus breeders to fingerling/smolt producers. The fry/fingerling producers raise the animals to a size suitable for the grow-out producers.

The high fecundity of fish can be fully utilised when producing grow-out animals since these will be slaughtered and inbreeding is not an issue as it is in the breeding nucleus. This makes the expense of maintaining broodstock low and facilitates the concentration of resources into one or a few multiplier units. The genetic gain obtained in the nucleus may in this way be effectively disseminated throughout the industry with a minimum

time lag. How this should be done most efficiently will vary with species and needs further study.

Significant non-additive genetic variation has been detected within fish populations (Rye and Mao, 1998; Pante et al., 2002). So that grow-out animals gain the maximum benefit from these non-additive effects, a systematic exploitation of non-additive genetic variation may be possible by repeated mating of progeny tested broodstock pairs or by development of specialized parent lines through some type of recurrent reciprocal selection (e.g. Toro, 1998).

Currently, all grow-out farmers are offered animals with similar genetic material that represents more or less the average genetic level in the nucleus. The potential for producing a more customised genetic material suited for specific farm environments and market segments may be possible. Grow-out animals could for example be produced from the best nucleus breeders for a single or few traits on request (in contrast with the nucleus where selection is for the overall aggregate genotype), preferably from progeny tested (see below) male and female breeders with highly accurate breeding values. However, the number of top quality breeders from the nucleus and consequently the number of grow-out animals may be limited. If early survival can be improved in the highly prolific marine species, a small number of breeders may still supply a large-scale commercial operation. Another alternative could be to establish separate selection lines outside the breeding nucleus. Each line could be reproduced using top male breeders from the nucleus with highly accurate breeding values for the desired trait(s) and female breeder from the line using individual selection. Individual selection of both male and female breeders within the line is also a possibility, but would be restricted to traits that can be recorded on the live breeding candidates. No tagging of animals would then be required in the line and several lines could therefore be maintained at a relatively low cost (e.g. by keeping several lines, tagged with fin-clipping, in the same rearing unit). After a few generations of selection the line(s) with highest market potential could be multiplied.

In addition to capitalising on additive and possible non-additive genetic effect at the multiplier level, complementary techniques like molecular genetics, sex and chromosome manipulations (Hulata, 2002) may be added to produce grow-out animals with the best performance. These techniques would need to be accommodated in the framework of the current classical selection programs.

12.12 Concluding remarks

Even if remarkable improvements of additive genetic performance have been obtained in some fish species, the application of livestock selection theory to farmed fish is still in its early stages. The opportunities for continued progress are good, but further investigations are needed so that genetic improvement programs can be developed that are optimal both with respect to genetic gain and costs.

Recent and future developments in molecular genetics present the possibility of using

genotype selection for performance traits (either through marker assisted or direct selection on quantitative trait loci or QTL's if this information is available). The optimal use of this information may require modification in the present mating and breeding design. The same applies to present technologies like sex and chromosome manipulations (Hulata, 2002). Methodology for efficiently integrating these technologies into classical breeding programs needs further investigation.

13. PREDICTION OF BREEDING VALUES

BJARNE GJERDE

13.1 Introduction

In selective breeding programs the main objective is to move the mean value of traits of economical importance in the desired direction. This change in the population mean due to selection is termed genetic gain (see Chapter 12). Genetic gain is obtained by selecting as parents for the next generation individuals that have a higher probability than other individuals in the population for transferring favourable values or genes to their offspring. This chapter deals with procedures to detect the best breeding individuals as parents for the next generation.

Parents pass on their genes and not their genotypes to their offspring. New genotypes are being created in each generation, as a result of coming together of new combinations of genes from the parents. Therefore the parents for the next generation should not be selected due to their genotypic value. A new measure of value is needed which will refer to genes and not to genotypes. This will enable us to assign a 'breeding value' to individuals, a value associated with the genes carried by the individual and transmitted to the offspring.

This new measure is the *'average effect of a gene'* which is the mean deviation from the population mean of individuals which received that gene from one parent, the gene received from the other parent having come at random from the population (Falconer and McKay, 1996). The breeding value of an individual for a given trait is equal to the sum of the average effects of the genes, the summation being made over the pair of alleles at each locus affecting the trait. Therefore, the average effects of the parents' genes determine the mean genotypic value of its progeny. Thus the *breeding value* of an individual can be defined as the value of the individual judged by the mean value of its progeny (Falconer and McKay, 1996). The breeding value for an individual, unlike average effect of a gene, can therefore be predicted based on phenotypic values recorded on its offspring. If an individual is mated to a number of individuals taken at random from the population, then its breeding value is twice the mean deviation of the progeny from the population mean. The deviation has to be doubled because the parent in question provides only half the genes in the progeny, the other half coming from the mates selected at random from the population.

Just as the average effect is a property of the gene and the population, so is the breeding value a property of the individual and the population from which its mates are drawn. Therefore when referring to an individual's breeding value, the population in which it was or will be mated must also be specified. Systematic and stochastic environmental effects and non-additive genetic effects will to a greater or lesser extent mask the true breeding value. Consequently, the true breeding value cannot be measured with 100% accuracy.

In animal breeding programs all traits of economic importance should be selected for simultaneously. This is more efficient than selecting for (i) one trait at a time until it is

T. Gjedrem (ed.), Selection and Breeding Programs in Aquaculture, 197–231.
© 2005 *Springer. Printed in the Netherlands.*

improved to the desired level (tandem selection) or (ii) assigning a threshold value for each trait below which all individuals are discarded, regardless of the superiority or inferiority of their other traits (independent culling). Selection based on an overall breeding value for *n* equally important and independent traits is $\sqrt{n}$ times as efficient as selection for the same traits, one at a time (Hazel and Lush, 1942). However, the genetic gain in any one trait by selection based on an overall breeding (index) value for n equally important and independent (uncorrelated) traits is only $1/\sqrt{n}$ times the gain obtained if selection was directed for that trait alone.

Genetic gain is directly proportional to the correlation ($r_{A\hat{A}}$) between the true and predicted breeding value (see Chapter 12); i.e., the accuracy of the predicted breeding value. Therefore accurate prediction of breeding values plays an important role in genetic improvement programmes. This is particularly the case for species in which relatively few individuals, through a high reproductive rate, have major influences on the genetic merit of the population.

Different types of methods are available for determining which individuals to be used as parents for the next generation (see Chapter 16). The methods differ with respect to type of relatives that provide information used for the selection decisions; the breeding candidate itself or relatives like full-sibs, half-sibs, offspring, parents etc. The objective of all these selection methods is to maximise the probability of correctly ranking the animals with respect to their breeding values. However, the magnitude of ($r_{A\hat{A}}$) may be very different for the different selection methods (see Chapter 11).

For each selection method, different genetic evaluation procedures can be applied to obtain the predicted breeding value. The most commonly used is the Selection Index Procedure (SIP) first applied in plant selection programmes by Smith (1936) and later developed for farm animals by Hazel (1943). Until the early 1970's, this was the standard procedure for the prediction of breeding values in animal breeding.

SIP is described in detail in many animal-breeding textbooks (e.g. Falconer and McKay, 1996; Cameron, 1997) and will therefore only be described briefly here for the sake of completeness. When selection is practised for one trait only and the available information is recorded on the breeding candidate and/or a limited number of relatives, SIP can be illustrated in a very simple manner. In section 13.3 this will be illustrated in detail for different selection methods and combination of methods.

An alternative and more powerful procedure is Best Linear Unbiased Prediction (BLUP, Henderson, 1984). In animal breeding this procedure was taken into use in the early 1970's and is now considered the state of the art for genetic evaluation of all animal and plant species. As opposed to SIP, BLUP provides simultaneous estimation of fixed (e.g. environmental) and random animal effects (e.g. breeding values). In addition, additive genetic relationships among all animals in the population can be incorporated and utilised more easily than with SIP. The derivation and properties of BLUP will be

outlined in section 13.5 while its application to fish will be described with an example in section 13.7.

13.2 Derivation of predictors

Now consider the simplest case for the prediction of breeding value of an individual. That is when the only information available is the phenotypic values of the individuals in a population for a single trait. Assuming no covariance between genetic and environmental effects, the phenotypic value can be partitioned into components attributable to different causes:

$$P_i = G_i + E_i = A_i + D_i + I_i + E_i \tag{13.1}$$

where

P_i =the phenotypic value of the individual

G_i =the genotypic value of the individual

A_i =the true breeding value of the individual

D_i =the effect due to intra locus gene interaction (dominance)

I_i = the effect due to inter locus gene interaction (epistasis)

E_i =the non-genetic effects that influence the phenotypic value

Thus in the simplest case the problem is how to obtain the best estimate of A_i when P_i is known. This is a prediction problem because the breeding value is a random variable that has not been realised in the population.

13.2.1 BEST PREDICTOR (BP)

To obtain a Best Predictor (BP) (smallest mean squares error and unbiased) requires that the form of the distribution of the observations (P_i) is known as well as all the moments (mean, variance, skewness, etc.). The BP is the conditional mean of A_i given P_i :

$$E(A_i | P_i) = \mu_A + A_i^{'} = A_i \tag{13.2}$$

where μ_A is the true overall mean breeding value and $A_i^{'}$ is the true breeding value for the ith individual as a deviation from μ_A. The mean squares error of BP is the sum of the variances of the predictor and the bias:

$$MSE = E\left[\left(A_i - \hat{A}_i\right)^2\right] = V(A_i) + V\left(A_i - \hat{A}_i\right) = V(A_i) \tag{13.3}$$

The bias is zero because the predictor is unbiased and thus this predictor of A_i has the smallest mean squares error of all predictors of A_i (Cochran, 1951). The form of this predictor depends on the distribution of P_i. The form could be either linear or non-linear functions of P_i.

13.2.2 BEST LINEAR PREDICTOR (BLP)

Restricting the predictor to be linear, then only the first (mean) and the second (variance) moments must be known. Thus we ignore or do not assume anything about the form of the distributions of the observations (P_i). Therefore the Best Linear Predictor (BLP) of A_i is also the conditional mean, but is a linear function of the observations (P_i):

$$E(A_i|P_i) = \hat{A}_i = \mu_A + \frac{\sigma_{AP}}{\sigma_P^2}(P_i - \mu_P) \qquad (13.4)$$

Assuming that the first and second moments are known ($\mu_A, \mu_P, \sigma_{AP}, \sigma_P^2$), BLP has the following properties: (a) minimise the mean squares error, (b) maximise the correlation between the true and predicted breeding value ($r_{A\hat{A}}$) and (c) maximise the genetic gain ($\Delta G = i r_{A\hat{A}} \sigma_{\hat{A}}$) by truncation selection on $\hat{A}_i$.

Assuming that P_i and A_i both are normally distributed with first and second moments known and that the predictor should be linear, BLP is also (d) unbiased and thus a best, linear, unbiased, predictor (BLUP) and (e) maximise the probability of correct ranking of the individuals with respect to $\hat{A}_i$.

This predictor (BLP) is identical to the predictor obtained by the selection index procedures (SIP) that are explained shortly below.

For some distributions of the observations, BP and BLP could be the same, for example for a normal distribution.

13.2.3 BEST LINEAR UNBIASED PREDICTOR (BLUP)

To assume that all means and variances are known, as required to obtain BLP, is not realistic in practical animal breeding. To do any prediction of random variables, the variances of the population parameters need to be known. Thus, BLUP assumes known variances, but allows the solution of fixed effects (means) to be unknown. In doing so the solutions for fixed effects must be unbiased. This can be obtained by applying a LaGrange Multiplier in deriving the predictor. The appropriate means are obtained as the Generalized Least Squares (GLS) solution for the fixed effect. Thus, BLUP is obtained as:

$$E(A_i|P_i) = \hat{A}_i = \hat{\mu}_A + \frac{\sigma_{AP}}{\sigma_P^2}(P_i - \hat{\mu}_P) \quad (13.5)$$

This equation is identical to the equation for BLP, except that the parameter estimates of the means ($\hat{\mu}_A, \hat{\mu}_P$) are used in place of the true values. The predicted breeding value is BLUP only if the (co)variances (σ_{AP}, σ_P^2) are known. In practice, the variances (and the means obtained as the GLS solution) are simultaneously estimated and used as though they were the true values, giving approximate or empirical BLUP (EBLUP). But BLUP is not necessarily the best predictor among all possible types of predictors, depending on the distribution of the random variables.

This section dealt only with the simplest case for prediction of breeding value; i.e., when the only information available is the phenotypic values of the individuals in a population for a single trait. Thus there is need for more general procedures where all types of information can be used in the selection decisions.

13.3 Selection Index Procedures (SIP)

SIP provides a reasonable solution to the problem of genetic evaluation of animals in a rather restricted situation; namely, when the animals and traits selected for have the same means or when the first moment (the means) and the second moments (variances) of the distribution of the observations is known. In practical animal breeding this is rarely the case. Therefore, when applying SIP the observed records are adjusted for the actual fixed effects (e.g. age, pond, sex, etc.) and assume that the means thus obtained are the true means.

13.3.1 SELECTION FOR SEVERAL TRAITS

SIP combines all available information about one as well as several traits recorded on the breeding candidate and/or its relatives; i.e., the selection criteria (P_i'), into an overall breeding (selection) objective or genetic merit (H). When *n* traits are included in the breeding objective, each with true breeding value A_i' and economic value a_i, then:

$$H = a_1A_1' + a_2A_2' + ... + a_nA_n' \quad (13.6)$$

where A_i' is expressed as a deviation from the true population mean, $A_i' = A_i - \mu_{A_i}$. The economic value of a trait is defined as the amount by which profit may be expected to increase for each unit of improvement in that trait (Hazel, 1943).

When *m* selection criteria are recorded, an estimate of H, denoted I, may be obtained as linear combinations of the criteria:

$$\hat{H} = I = b_1P_1' + b_2P_2' + ... + b_mP_m' \quad (13.7)$$

where the $P_i^{'}$'s are the adjusted record expressed as a deviation from the adjusted population mean, $P_i^{'} = P_i - \mu_{P_i}$, and the b_i's are weighting factors obtained by solving a set of equations such that the correlation between the true (H) and estimated (I) breeding objective is maximised. The number of selection criteria (m) may be different (equal, larger or smaller) than the number of traits (n) in the breeding objective.

Thus the use of SIP for the prediction of breeding values involves a stepwise procedure:

1. Appropriate adjustment of the recorded selection criteria for systematic environmental effects
2. Find the set of weighting factors, b_i, that maximise the correlation between the true and estimated breeding value (r_{HI}). This is obtained simultaneously for all b_i by solving a set of normal equations, which in matrix notation can be written as $\mathbf{Pb} = \mathbf{Aa}$ and $\mathbf{b} = \mathbf{P}^{-1}\mathbf{Aa}$, where $\mathbf{P}$ is the phenotypic variance covariance matrix among the adjusted selection criteria and with inverse $\mathbf{P}^{-1}$; $\mathbf{A}$ is the additive genetic variance and covariance matrix among the traits in the selection objective; $\mathbf{a}$ is a vector of economic values; and $\mathbf{b}$ is a vector of unknown partial regression coefficients or weighting factors to be estimated
3. The predicted breeding value for each animal (I) is found by applying equation (2) with the adjusted observations ($P_i^{'}$) obtained from step 1 and the b_i's obtained from step 2

The estimated breeding values thus obtained are expressed as deviations from the population mean of the overall breeding objective (μ_H). The predicted breeding values thus obtained are BLUP only if all the records have the same mean (the overall mean is the only fixed effect), the means of the fixed effects are known or the means are the Generalized Least Squares (GLS) solution of the fixed effects.

When solving for $\mathbf{b}$ in the normal equations, $\mathbf{Pb} = \mathbf{Aa}$, a restriction is imposed such that the regression of true on predicted breeding value is unity ($\frac{\sigma_{HI}}{\sigma_I^2} = 1$). Thus the variance of the predicted breeding values (σ_I^2) is equal to its covariance with the true breeding value (σ_{HI}). This implies that the regression coefficients (b_i) are scaled in such a way that one unit of the predicted breeding value is equivalent to one unit of the true breeding value.

13.3.2 SELECTION FOR ONE TRAIT

When selection is practised for one trait only and the available information is recorded on the breeding candidate and/or a given number of relatives (e.g., full-, half-sibs and/or progeny), the use of SIP can be illustrated with the necessary variances and covariances expressed as a function of the full-sib family size (n) and intraclass correlation, $t = rh^2 + c^2$, where r is the additive genetic relationship among full-sibs (r=0.5) or half-sibs

(r=0.25), h^2 is the heritability, and c^2 is an effect common to full-sibs other than additive genetics (see e.g. Falconer and McKay, 1996; Cameron, 1997).

Furthermore it is assume that sires, dams and sires and dams are not related and that a mating design with dams nested within sires (m = number of dams per sire) is applied. Then the following set of parameters can be defined:

$\sigma_P^2 = \sigma_s^2 + \sigma_d^2 + \sigma_w^2 + \sigma_c^2 + \sigma_e^2$; the phenotypic variance,

$\sigma_A^2 = h^2\sigma_P^2$; the additive genetic variance = $\sigma_s^2 + \sigma_d^2 + \sigma_w^2$,

h^2 = the heritability of the trait,

$\sigma_s^2 = 0.25\sigma_A^2$; the additive genetic variance among sires,

$\sigma_d^2 = 0.25\sigma_A^2$; the additive genetic variance among dams within sires,

$\sigma_w^2 = 0.50\sigma_A^2$; the additive genetic variance within full-sibs,

$\sigma_c^2 = c^2\sigma_P^2$; the variance common to full-sibs other than additive genetics,

c^2 = the proportion of the variance common to full-sibs other than additive genetics,

$\sigma_e^2 = \sigma_P^2 - \sigma_A^2 - \sigma_c^2$; the random environmental variance,

The following quantities can be derived:

$\sigma_{\overline{P}_f}^2 = \sigma_s^2 + \sigma_d^2 + \sigma_c^2 + (\sigma_w^2 + \sigma_e^2)/n$; i.e., the phenotypic variance of a full-sib family mean,

$\sigma_{P_{w1}}^2 = \sigma_w^2 + \sigma_e^2 - (\sigma_w^2 + \sigma_e^2)/n$; i.e., the phenotypic variance of within full-sib family deviations $P_{w1} = P_i - \overline{P}_f$ when the individual is included in the full-sib mean, and

$\sigma_{P_{w2}}^2 = \sigma_w^2 + \sigma_e^2 + (\sigma_w^2 + \sigma_e^2)/n$; i.e., the phenotypic variance of within full-sib family deviations $P_{w2} = P_i - \overline{P}_f$ when the individual is not included in the full-sib mean.

13.3.3 INDIVIDUAL SELECTION

When only one trait is included in the selection objective ($H = a_1A_1 = A_1$) and the only selection criterion is one observation of this trait recorded on the breeding candidates (the phenotypic value), then the breeding value of an individual can be calculated as:

$$\hat{A}_i = \hat{\mu}_A + \frac{\sigma_{AP}}{\sigma_P^2}(P_i - \hat{\mu}_P) = \hat{\mu}_A + h^2(P_i - \hat{\mu}_P) \qquad (13.8)$$

The weighting factor is the regression of the estimated breeding value on the phenotypic value, which by definition is equal to the heritability of the trait. This assumes that the covariances of additive genetic effects with non-additive genetic and environmental effects are zero.

13.3.4 WITHIN FULL-SIB FAMILY SELECTION

The criterion of selection for within full-sib family selection is the deviation of each individual from the phenotypic mean value of the full-sib family it belongs. The prediction equation may be written as:

$$\hat{A}_i = \hat{\mu}_A + \frac{\sigma_{AP_w}}{\sigma^2_{P_w}} P_w = \hat{\mu}_A + \frac{\sigma_{AP_w}}{\sigma^2_{P_w}} (P_i - \overline{P}_f) \qquad (13.9)$$

The weighting factor is the regression of the true breeding values of the breeding candidates on the deviations from the full-sib family mean $P_w = P_i - \overline{P}_f$. When the phenotypic value of the individual is included in the full-sib mean, $\sigma^2_{P_{w1}} = \sigma^2_w + \sigma^2_e - (\sigma^2_w + \sigma^2_e)/n$, the weighting factor is $\frac{\sigma^2_w - \sigma^2_w/n}{\sigma^2_w + \sigma^2_e}$. When n increases (large families), the weighting factor approaches h^2_w, i.e. the within full-sib family heritability. When the phenotypic value of the individual is not included in the full-sib mean, $\sigma^2_{P_{w2}} = \sigma^2_w + \sigma^2_e + (\sigma^2_w + \sigma^2_e)/n$, the weighting factor is $\sigma^2_w / (\sigma^2_w + \sigma^2_e) = h^2_w$, i.e. the within full-sib family heritability.

13.3.5 SIB SELECTION

When the trait selected for is not recorded on the breeding candidates, only on its full-sibs, the estimated breeding value of the breeding candidate can be calculated as:

$$\hat{A}_i = \hat{\mu}_A + \frac{\sigma_{A\overline{P}_f}}{\sigma^2_{\overline{P}_f}} (\overline{P}_f - \hat{\mu}_P) \qquad (13.10)$$

$\overline{P}_f$ is the phenotypic mean of the measured full-sibs and the weighting factor is the regression of the true breeding values of the breeding candidates on the mean phenotypic value of their measured full-sibs. The weighting factor may be written as:

$$\frac{\sigma_{A\overline{P}_f}}{\sigma^2_{\overline{P}_f}} = \frac{\sigma^2_s + \sigma^2_d}{\sigma^2_s + \sigma^2_d + \sigma^2_c + (\sigma^2_w + \sigma^2_e)/n} \qquad (13.11)$$

When the size of the full-sib family (*n*) increases, the within full-sib additive genetic and random environmental variances approaches zero and the denominator of the weighting factor approaches $\sigma^2_s + \sigma^2_d + \sigma^2_c$. For $\hat{\sigma}^2_c = 0$, the latter quantity is equal to the additive genetic variance of full-sib family means, $\sigma^2_s + \sigma^2_d = 0.5\sigma^2_a$, and the weighting factor is unity. Thus under these circumstance, large n and $\sigma^2_c = 0$, the phenotypic mean of the full-sib family is equal to its additive genetic mean and thus not regressed towards the population mean.

13.3.6 FAMILY SELECTION

When the breeding candidates are measured and included in the full-sib family means, but are only used to determine the family mean, the term family selection is used. Thus under family selection the within family deviations are given zero weight. Under these circumstances the numerator of the weighting factor is $\sigma_{A\bar{P}_f} = \sigma_s^2 + \sigma_d^2 + \sigma_w^2 / n$ and approaches $\sigma_s^2 + \sigma_d^2 = 0.5\sigma_a^2$ when n increases. The denominator is as for sib selection. Thus for large full-sib families, sib and family selection is equal.

13.3.7 OWN, FULL- AND HALF-SIB SELECTION

When the trait selected for is recorded both on the breeding candidate and its full- and half-sibs, the prediction equation will contain more than one term. In fish where a nested mating design is frequently used with one male mated to two or more females or vice versa (see Chapter 12), there are three different selection criteria or sources of information; the individual, its full-sibs and its half-sibs, resulting in a prediction equation with three terms. However, the group of half-sibs is composed of two (two dams per sire) or more (three or more dams per sire) full-sib groups. In this case it is convenient that the prediction equation includes, in addition to a term for the breeding candidate, a term for each full-sib family:

$$\hat{A}_i = \hat{\mu}_A + b_i(P_i - \mu_P) + b_f(\bar{P}_f - \mu_P) + b_{h_1}(\bar{P}_{h_1} - \mu_P) + b_{h_2}(\bar{P}_{h_2} - \mu_P) + \ldots\ldots\ldots + b_{m-1}(\bar{P}_{h_{m-1}} - \mu_P) \quad (13.12)$$

where $\bar{P}_f$ is the mean phenotypic value of the measured full-sibs including the breeding candidate; $\bar{P}_{h_1}, \bar{P}_{h_2}, \ldots, \bar{P}_{h_{m-1}}$ are the mean phenotypic values of each of the additional groups of full-sibs (i.e. half-sibs to the breeding candidate); $b_f, b_{h_1}, b_{h_2}, \ldots, b_{h_{m-1}}$ are the partial regression coefficients or weighting factors to be estimated and m is the number of females mated to each male.

Given a design with one male mated to three females (m = 3), the matrices **P** and **A** of the normal equations **Pb** = **Aa** are as follows (for one trait in the breeding objective, i.e., **a** =1):

$$\mathbf{P} = \begin{bmatrix} \sigma_P^2 & \sigma_s^2 + \sigma_d^2 + \sigma_c^2 + (\sigma_w^2 + \sigma_e^2)/n_1 & \sigma_s^2 & \sigma_s^2 \\ \text{sym} & \sigma_s^2 + \sigma_d^2 + \sigma_c^2 + (\sigma_w^2 + \sigma_e^2)/n_1 & \sigma_s^2 & \sigma_s^2 \\ \text{sym} & \text{sym} & \sigma_s^2 + \sigma_d^2 + \sigma_c^2 + (\sigma_w^2 + \sigma_e^2)/n_2 & \sigma_s^2 \\ \text{sym} & \text{sym} & \text{sym} & \sigma_s^2 + \sigma_d^2 + \sigma_c^2 + (\sigma_w^2 + \sigma_e^2)/n_3 \end{bmatrix}$$

13.13)

where n_1, n_2 and n_3 is the number of recorded individuals for each of the three full-sib family groups.

$$\mathbf{A} = \begin{bmatrix} \sigma_a^2 \\ \sigma_s^2 + \sigma_d^2 + \sigma_w^2 / n_1 \\ \sigma_s^2 \\ \sigma_s^2 \end{bmatrix} \tag{13.14}$$

For each additional full-sib family, one additional row and column must be added to the matrix $\mathbf{P}$ and one additional row to the matrix $\mathbf{A}$. For paired mating (each male mated to only one female), $\mathbf{P}$ will contain the two first rows and columns of $\mathbf{P}$ and $\mathbf{A}$ and the two first rows of $\mathbf{A}$. If the phenotypic value of the breeding candidate is not available, the first row and column of $\mathbf{P}$ and the first row of $\mathbf{A}$ has to be deleted.

When the observation on the breeding candidate is not included in the phenotypic mean of its full-sibs, both the phenotypic covariance between the individual and the mean of its full-sibs is $\sigma_s^2 + \sigma_d^2 + \sigma_c^2$ instead of $\sigma_s^2 + \sigma_d^2 + \sigma_c^2 + (\sigma_w^2 + \sigma_e^2)/n$ and the genetic covariance is $\sigma_s^2 + \sigma_d^2 + \sigma_w^2/n$. This requires that a separate full-sib mean is calculated for each individual while only one mean for each full-sib group is required when the individual is included in the full-sib mean.

For both approaches, the weighting factors are obtained as $\mathbf{b} = \mathbf{P}^{-1}\mathbf{Aa}$. Thus, a separate $\mathbf{P}^{-1}$ is required for each available combination of n_1, n_2, n_3. With access to a computer-software that invert matrices this is not a problem.

Nevertheless, to facilitate computation the above prediction equation could be set up using independent linear functions of the records so that $\mathbf{P}$ becomes diagonal (Wray and Hill, 1989). In this case the predicted breeding value may be obtained as:

$$\hat{A}_i = \hat{\mu}_A + b_i(P_i - \overline{P}_f) + b_f(\overline{P}_f - \overline{P}_h) + b_h(\overline{P}_h - \mu_P) \tag{13.15}$$

when a sire is mated to two or more dams, and without the last part if paired mating is practised; i.e., one sire per dam. $\overline{P}_h$ is the phenotypic (least-square) mean of the half-sibs and $\overline{P}_f$ is included in $\overline{P}_h$.

Therefore, for a mating design with two dams per sire:

$$\mathbf{P} = \begin{bmatrix} \sigma_w^2 + \sigma_e^2 & 0 & 0 \\ 0 & \sigma_d^2 + \sigma_c^2 + (\sigma_w^2 + \sigma_e^2)/n_1 & 0 \\ 0 & 0 & \sigma_s^2 + (\sigma_d^2 + \sigma_c^2)/m + (\sigma_w^2 + \sigma_e^2)/(n_1 + n_2) \end{bmatrix} \tag{13.16}$$

$$\mathbf{A} = \begin{bmatrix} \sigma_w^2 \\ \sigma_d^2 + \sigma_w^2/n_1 \\ \sigma_s^2 + \sigma_d^2/m + \sigma_w^2/(n_1 + n_2) \end{bmatrix} \tag{13.17}$$

Thus, since $\mathbf{P}$ is diagonal, its inverse can simply be obtained as the inverse of each of the diagonal elements and the weighting factors calculated independently of each other as:

$$b_i = \frac{\sigma_w^2}{\sigma_w^2 + \sigma_e^2} \tag{13.18}$$

$$b_f = \frac{\sigma_d^2 + \sigma_w^2/n_1}{\sigma_d^2 + \sigma_c^2 + (\sigma_w^2 + \sigma_e^2)/n_1} \tag{13.19}$$

$$b_h = \frac{\sigma_s^2 + \sigma_d^2/m + \sigma_w^2/(n_1 + n_2)}{\sigma_s^2 + (\sigma_d^2 + \sigma_c^2)/m + (\sigma_w^2 + \sigma_e^2)/(n_1 + n_2)} \tag{13.20}$$

Worthy of note is that individual (mass) selection is a special case of this index where $b_i = b_f = b_h = 1$ (Villenueva and Woolliams, 1997).

For balanced mating designs, i.e. equal full-sib family sizes ($n_1 + n_2 + ..n_m = mn$) and equal number of dams per sire, the above predictions gives BLUP-breeding values if the overall mean is the only fixed effect in the model. For unbalanced design the discrepancy from BLUP will be marginal for large full-sib family sizes (n > 5).

13.3.8 PROGENY TESTING

The criterion of selection for progeny testing is the mean value of an individual's progeny and thus the operational definition of breeding value. Consider an example where we want to predict the breeding value of a sire that is mated to three different dams. The selection criterion is a trait recorded on the offspring of the three dams (full-sib families). If we assume that neither of the parents is measured, the prediction equation is:

$$\hat{A}_i = \hat{\mu}_A + b_{f_1}(\overline{P}_{f_1} - \mu_P) + b_{f_2}(\overline{P}_{f_2} - \mu_P) + b_{f_3}(\overline{P}_{f_3} - \mu_P) \tag{13.21}$$

where $\overline{P}_{f_1}, \overline{P}_{f_2}, \overline{P}_{f_3}$ are the mean phenotypic value of each of the three full-sib families.

The weighting factors $b_{f_1}, b_{f_2}, b_{f_3}$ are obtained as $\mathbf{b} = \mathbf{P}^{-1}\mathbf{A}\mathbf{a}$ where

$$\mathbf{P} = \begin{bmatrix} \sigma_s^2 + \sigma_d^2 + \sigma_c^2 + (\sigma_w^2 + \sigma_e^2)/n_1 & \sigma_s^2 & \sigma_s^2 \\ \sigma_s^2 & \sigma_s^2 + \sigma_d^2 + \sigma_c^2 + (\sigma_w^2 + \sigma_e^2)/n_2 & \sigma_s^2 \\ \sigma_s^2 & \sigma_s^2 & \sigma_s^2 + \sigma_d^2 + \sigma_c^2 + (\sigma_w^2 + \sigma_e^2)/n_3 \end{bmatrix} \tag{13.22}$$

$$\mathbf{A} = \begin{bmatrix} \sigma_s^2 + \sigma_d^2 \\ \sigma_s^2 + \sigma_d^2 \\ \sigma_s^2 + \sigma_d^2 \end{bmatrix} \tag{13.23}$$

If the objective is to calculate the breeding value of one of the dams, the prediction equation is:

$$\hat{A}_i = \hat{\mu}_A + \frac{\hat{\sigma}_{A\bar{P}_f}}{\hat{\sigma}^2_{\bar{P}_f}}(\bar{P}_f - \hat{\mu}_P) \tag{13.24}$$

where the weighting factor is the regression of the true breeding value of the dam on the mean phenotypic performance of her full-sib offspring. Thus the rows and columns of the above matrices **P** and **A** pertaining to the two other dams have to be deleted. Thus **P** and **A** contain only one element each and the weighting factor is:

$$\frac{\hat{\sigma}_{A\bar{P}_f}}{\hat{\sigma}^2_{\bar{P}_f}} = \frac{\hat{\sigma}_s^2 + \hat{\sigma}_d^2}{\hat{\sigma}_s^2 + \hat{\sigma}_d^2 + \hat{\sigma}_c^2 + (\hat{\sigma}_w^2 + \hat{\sigma}_e^2)/n} \tag{13.25}$$

and thus equal to the weighting factor for sib selection. Therefore, in a single paired mating design and with no record on the parents, the predicted breeding value of the dam (or the sire) is equal to the predicted breeding value of a randomly chosen offspring.

13.4 Accuracy of predicted breeding value

The accuracy of the predicted breeding value (i.e., the accuracy of selection) is defined as the correlation between the true and predicted breeding values ($r_{A\hat{A}}$) and is a measure of how well the breeding candidates are ranked with respect to their true breeding value. The accuracy squared ($r^2_{A\hat{A}}$) expresses the proportion of the true additive genetic variance for a trait or combination of traits (breeding objective) that is accounted for by the predicted breeding value; conversely $1 - r^2_{A\hat{A}}$ is the proportion not accounted for (see next section).

By definition $\sigma_{A\hat{A}} = \sigma^2_{\hat{A}}$ (see Chapter 9). Therefore, the correlation between the true (A) and predicted ($\hat{A}$) breeding value is equal to the ratio between the standard deviation of the predicted and the true breeding value:

$$r_{A\hat{A}} = \frac{\sigma_{A\hat{A}}}{\sigma_A \sigma_{\hat{A}}} = \frac{\sigma^2_{\hat{A}}}{\sigma_A \sigma_{\hat{A}}} = \frac{\sigma_{\hat{A}}}{\sigma_A} \tag{13.26}$$

The accuracy is obtained by, first deriving the variance of the predicted breeding value ($\sigma^2_{\hat{A}}$), and secondly taking the square root of this variance $\sqrt{\sigma^2_{\hat{A}}} = \sigma_{\hat{A}}$.

The derivation of accuracy of predicted breeding values will be shown for the following selection methods: individual selection, within family selection, sib selection and family selection and for combined (own, full- and half-sib) selection with independent sources of information.

13.4.1 INDIVIDUAL SELECTION

The variance of the predicted breeding values is:

$$\sigma^2_{\hat{A}} = \left[\frac{\sigma_{AP}}{\sigma^2_P}\right]^2 \sigma^2_P = (h^2)^2 \sigma^2_P \tag{13.27}$$

and the accuracy:

$$r_{A\hat{A}} = \frac{\sqrt{(h^2)^2 \sigma^2_P}}{\sigma_A} = \frac{h^2 \sigma_P}{\sigma_A} = h \tag{13.28}$$

Thus for phenotypic selection the accuracy of the predicted breeding value is the square root of the heritability.

13.4.2 WITHIN FAMILY SELECTION

When the individual is included in the full-sib family mean, the variance of the predicted breeding values is:

$$\sigma^2_{\hat{A}} = \left[\frac{\sigma_{AP_w}}{\sigma^2_{P_w}}\right]^2 \sigma^2_{P_{w1}} = \left[\frac{\sigma^2_w - \sigma^2_w/n}{\sigma^2_w + \sigma^2_e}\right]^2 (\sigma^2_w + \sigma^2_e)(1 - 1/n) \tag{13.29}$$

and the accuracy:

$$r_{A\hat{A}} = \frac{\sigma_{\hat{A}}}{\sigma_w} = \frac{(\sigma^2_w - \sigma^2_w/n)\sqrt{1 - 1/n}}{\sigma_w} \tag{13.30}$$

When the individual is not included in the full-sib family mean, the variance of the predicted breeding values is:

$$\sigma^2_{\hat{A}} = \left[\frac{\sigma_{AP_w}}{\sigma^2_{P_w}}\right]^2 \sigma^2_{P_{w2}} = \left[\frac{\sigma^2_w}{\sigma^2_w + \sigma^2_e}\right]^2 (\sigma^2_w + \sigma^2_e)(1 + 1/n) \tag{13.31}$$

and the accuracy:
$$r_{A\hat{A}} = \frac{\sigma_{\hat{A}}}{\sigma_w} = \frac{\sigma_w^2\sqrt{1+1/n}}{\sigma_w} \tag{13.32}$$

13.4.3 SIB SELECTION

The variance of the predicted breeding values is:

$$\sigma_{\hat{A}}^2 = \left[\frac{\sigma_{A\bar{P}_f}^2}{\sigma_{\bar{P}_f}^2}\right]^2 \sigma_{\bar{P}_f}^2 = \frac{(\sigma_{A\bar{P}_f}^2)^2}{\sigma_{\bar{P}_f}^2} = \frac{(\sigma_s^2+\sigma_d^2)^2}{\sigma_s^2+\sigma_d^2+\sigma_c^2+(\sigma_w^2+\sigma_e^2)/n} \tag{13.33}$$

Since the numerator is equal to $(0.5\sigma_A^2)^2$, the accuracy squared is:

$$r_{A\hat{A}}^2 = \frac{\sigma_{\hat{A}}^2}{\sigma_A^2} = \frac{0.5\sigma_w^2}{\sigma_s^2+\sigma_d^2+\sigma_c^2+(\sigma_w^2+\sigma_e^2)/n} \tag{13.34}$$

When n is large and $\sigma_c^2 = 0$ the accuracy attains a maximum value of $r_{A\hat{A}} = \sqrt{0.5} = 0.7071$ and thus equal to the square root of the additive genetic relationship among full-sibs.

13.4.4 FAMILY SELECTION

The variance of the predicted breeding values is:

$$\sigma_{\hat{A}}^2 = \left[\frac{\sigma_{A\bar{P}_f}}{\sigma_{\bar{P}_f}^2}\right]^2 \sigma_{\bar{P}_f}^2 = \frac{(\sigma_{A\bar{P}_f}^2)^2}{\sigma_{\bar{P}_f}^2} = \frac{(\sigma_s^2+\sigma_d^2+\sigma_w^2/n)^2}{\sigma_s^2+\sigma_d^2+\sigma_c^2+(\sigma_w^2+\sigma_e^2)/n} \tag{13.35}$$

and the accuracy squared:

$$r_{A\hat{A}}^2 = \frac{\sigma_{\hat{A}}^2}{\sigma_A^2} = \frac{(\sigma_s^2+\sigma_d^2+\sigma_w^2/n)^2}{\sigma_A^2(\sigma_s^2+\sigma_d^2+\sigma_c^2+(\sigma_w^2+\sigma_e^2)/n)} \tag{13.36}$$

When *n* is large the accuracy is equal to the accuracy for sib selection.

13.4.5 COMBINED OWN, FULL- AND HALF-SIB SELECTION

When the predicted breeding value is obtained using independent linear functions of the records (section 13.3.5), all co-variances among the different sources of information are

zero. Consequently, the variance of the predicted breeding value is equal to the sum of the variance of each element and is according to Wray et al. (1994):

$$\sigma^2_{\hat{A}} = b_i^2 k_i + b_f^2 k_f + b_h^2 k_h \tag{13.37}$$

where b_i, b_f, b_h are defined earlier (section 13.3.5) and

$$k_i = (\sigma^2_w + \sigma^2_e)(1 - 1/n) \tag{13.38}$$

$$k_f = \left[\sigma^2_d + \sigma^2_c + (1/n)(\sigma^2_w + \sigma^2_e)\right]\left[1 - 1/m\right] \tag{13.39}$$

$$k_h = \sigma^2_s + \left[\sigma^2_d + \sigma^2_c + (1/n)(\sigma^2_w + \sigma^2_e)\right]\left[1/m\right] \tag{13.40}$$

and the accuracy :

$$r_{A\hat{A}} = \frac{\sigma_{\hat{A}}}{\sigma_A} \tag{13.41}$$

13.4.6 ACCURACY FOR DIFFERENT SELECTION METHODS

Figure 13.1 (a-d) shows the accuracy of predicted breeding values for four different selection methods when selection is practised for a single trait with different heritabilities (h^2), proportions of variance common to full-sibs other than additive genetics (c^2) and sizes of the full-sib family (n). As expected, accuracy increases with increasing h^2 and decreases with increasing c^2. With respect to accuracy, the value of one additional sib record is larger for small than for large family sizes and larger for combined than for sib selection.

The accuracy for progeny testing is shown in Figure 13.2. When number of mates is equal to one (single pair mating) the accuracy is equal to the accuracy for sib selection. The accuracy increases with increasing number of mates and with size of the full-sib families. For high number of mates (> 6) and large full-sib famiies (> 50), the accuracy is very high (> 0.90) and very similar for h^2=0.10 and h^2=0.30.

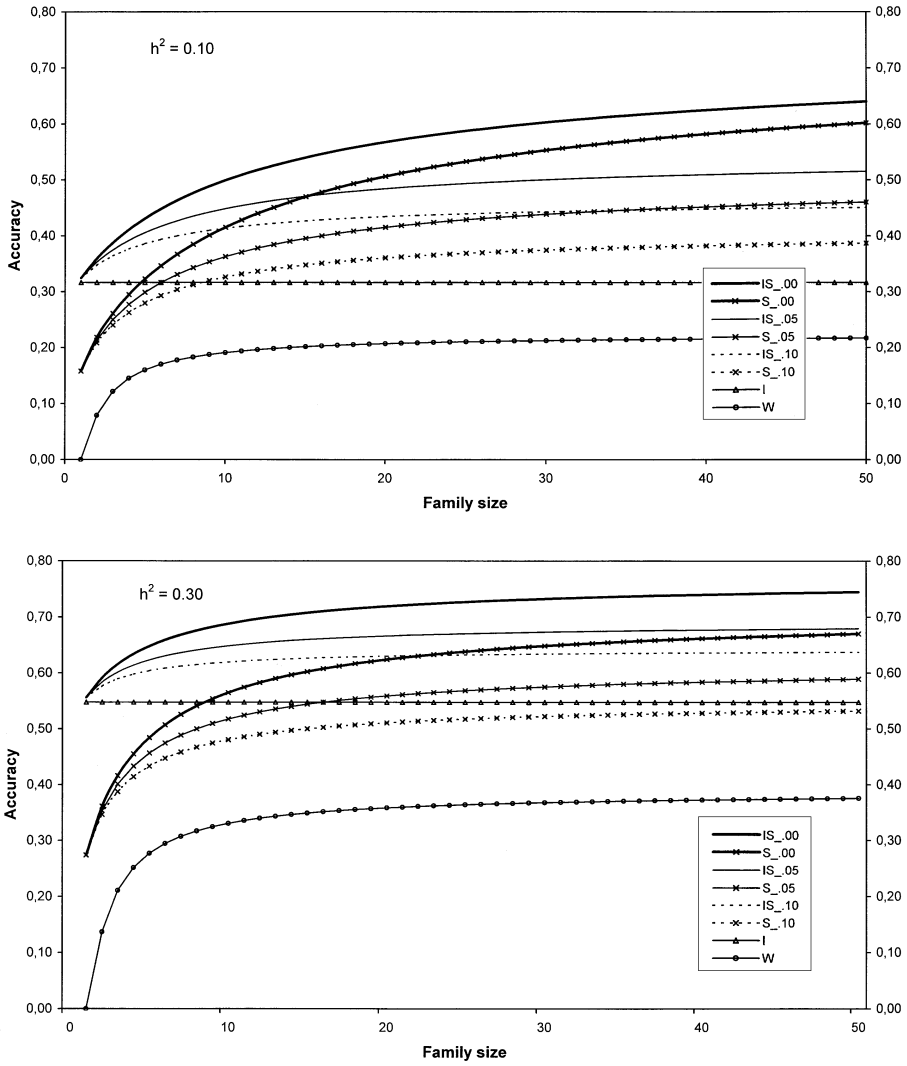


Figure 13.1 Accuracy of different selection methods as a function of full-sib family size for two different heritabilities (h^2) and three different magnitudes of the effect common to full-sibs other than additive genetics (c^2). I = Individual selection; W = within family selection; S_.00 = Sib selection, $c^2 = 0.00$; IS_.05 = Combined (own and sib information) selection, $c^2 = 0.05$.

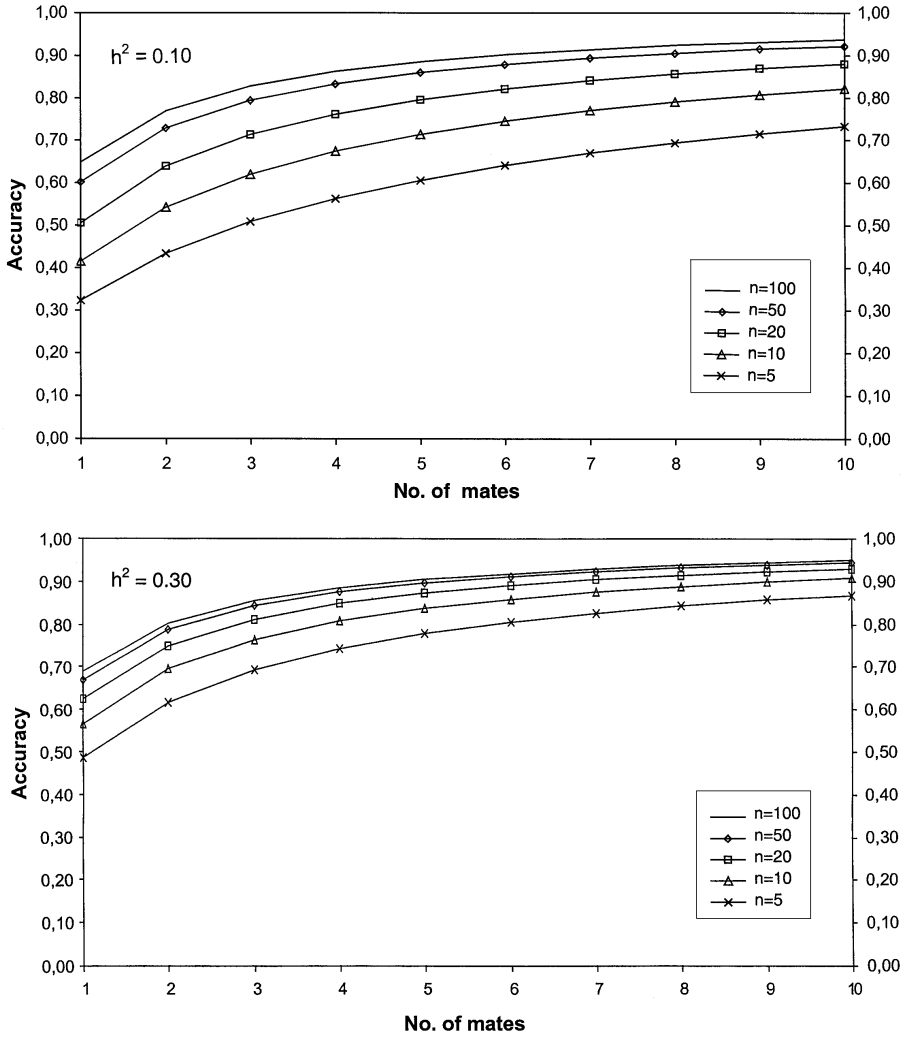


Figure 13.2 Accuracy of selection for progeny testing as a function of number of mates and size of full-sib family (n) for two different heritabilities.

13.4.7 PREDICTION ERROR VARIANCE

The accuracy of selection squared ($r^2_{A\hat{A}}$) expresses the proportion of the true additive genetic variance accounted for by the predicted breeding value. The prediction error variance (PEV) is the proportion of the true additive genetic variance not accounted for:

$$PEV = \left(1 - r^2_{A\hat{A}}\right)\sigma^2_A = \sigma^2_A - r^2_{A\hat{A}}\sigma^2_A = \sigma^2_A - \sigma^2_{\hat{A}} = \sigma^2_{(A-\hat{A})} \qquad (13.42)$$

Thus when the amount of information about the predicted breeding value increases, $r^2_{A\hat{A}}$ and therefore $\sigma^2_{\hat{A}}$ increases and the prediction error variance decreases. It follows that the true additive genetic variance is the sum of the variance of the predicted breeding values and the prediction error variance.

As an example, for individual selection the prediction error variance is:

$$PEV = \left(1 - h^2\right)\sigma^2_A \tag{13.43}$$

13.5 Best Linear Unbiased Predictor (BLUP)

Henderson (1949) developed a methodology called best linear unbiased prediction (BLUP) by which estimates of fixed effects (BLUE = Best Linear Unbiased Estimates) and prediction of breeding values (BLUP) are obtained simultaneously. However, it was not until the early 1970's that BLUP was applied to genetic evaluation of animals. BLUP has found widespread usage in genetic evaluation of domestic animals and also cultivated crop species because of its desirable properties and by the steady increase in computing power. The properties of BLUP are incorporated in the name:

- Best – it maximizes the correlation between the true (A) and predicted $(\hat{A})$ breeding value or minimizes prediction error variance, $PEV = \sigma^2_A - \sigma^2_{\hat{A}} = \sigma^2_{(A-\hat{A})}$
- Linear – the predictors are linear functions of the observations
- Unbiased – estimation of realized values of a random variable such as animal breeding values and of estimable functions of fixed effects are unbiased, $E(A|\hat{A}) = \hat{A}$
- Prediction – involves prediction of true breeding values

The theoretical background for BLUP will be briefly presented in the following. In section 13.4 the application of BLUP for the genetic evaluation of fish will be illustrated Consider the following equation for a mixed linear animal model:

$$\mathbf{y = Xb + Zu + e} \tag{13.44}$$

where

y = n x 1 vector of observations where n = number of records,

b = p x 1vector of unknown fixed effects where p = number of levels for fixed effects

u = q x 1vector of unknown random animal additive genetic effects where q =number of levels for random effect,

e = n x 1 vector of unknown random residual effects corresponding to the elements of **y** ,

$\mathbf{X}$ = a known design matrix of order n x p, which relates the elements of $\mathbf{y}$ to those of $\mathbf{b}$,

$\mathbf{Z}$ = a known design matrix of order n x q, which relates the elements of $\mathbf{y}$ to those of $\mathbf{u}$.

The fixed effect $\mathbf{b}$ may be partitioned into more set of factors depending on the situation.

The expectations are: $E(\mathbf{y}) = \mathbf{Xb}$; $E(\mathbf{u}) = \mathbf{0}$; $E(\mathbf{e}) = \mathbf{0}$. The residual effects, which include random effects other than additive genetics (e.g., environmental and non-additive genetic effects) are independently distributed with variance σ_e^2. Therefore, $\text{var}(\mathbf{e}) = \mathbf{I}\sigma_e^2 = \mathbf{R}$; $\text{var}(\mathbf{u}) = \mathbf{A}\sigma_a^2 = \mathbf{G}$ and $\text{cov}(\mathbf{u},\mathbf{e}) = \text{cov}(\mathbf{e},\mathbf{u}) = 0$, where $\mathbf{A}$ is the additive genetic relationship matrix. Then

$$\begin{aligned}\text{var}(\mathbf{y}) = \mathbf{V} &= \text{var}(\mathbf{Zu} + \mathbf{e}) \\ &= \mathbf{Z}\,\text{var}(\mathbf{u})\mathbf{Z}' + \text{var}(\mathbf{e}) + \text{cov}(\mathbf{Zu},\mathbf{e}) + \text{cov}(\mathbf{e},\mathbf{Zu}) \\ &= \mathbf{ZGZ}' + \mathbf{R} + \mathbf{Z}\,\text{cov}(\mathbf{u},\mathbf{e}) + \text{cov}(\mathbf{e},\mathbf{u})\mathbf{Z}' \\ &= \mathbf{ZGZ}' + \mathbf{R}\end{aligned} \tag{13.45}$$

The problem is to predict $\mathbf{K}'\mathbf{b} + \mathbf{M}'\mathbf{u}$ (predictand) by a function of **y** (predictor) where **K** and **M** are matrices specifying the contrasts of interest among the level of the fixed and random effects, respectively. The best predictor of $\mathbf{K}'\mathbf{b} + \mathbf{M}'\mathbf{u}$ is $\text{BP} = \mathbf{E}(\mathbf{K}'\mathbf{b} + \mathbf{M}'\mathbf{u} \,|\, \mathbf{y}) = \mathbf{K}'\mathbf{b} + \mathbf{M}'\mathbf{u}$.

By restricting the predictor to be

1. a linear function of y, $\mathbf{L}'\mathbf{y}$ (i.e. only the first and second moment need to be known),
2. estimable, i.e., $\mathbf{L}'\mathbf{Xb} = \mathbf{K}'\mathbf{b}$, that is $\mathbf{L}'\mathbf{X} - \mathbf{K}' = 0$ for all values of **b**,
3. minimum variance, i.e., $\text{PEV} = \text{var}(\mathbf{L}'\mathbf{y} - \mathbf{K}'\mathbf{b} - \mathbf{M}'\mathbf{u})$
4. unbiased, i.e., $E(\mathbf{L}'\mathbf{y}) = E(\mathbf{K}'\mathbf{b} + \mathbf{M}'\mathbf{u}) = \mathbf{K}'\mathbf{b} = \mathbf{L}'\mathbf{Xb}$,

then (see Henderson, 1973 and 1984 for detailed derivations) BLUP of $\mathbf{K}'\mathbf{b} + \mathbf{M}'\mathbf{u}$ is

$$\mathbf{L}'\mathbf{y} = \mathbf{K}'\hat{\mathbf{b}} + \mathbf{M}'\mathbf{GZ}'\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\hat{\mathbf{b}}) \tag{13.46}$$

where

$$\hat{\mathbf{b}} = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-}\mathbf{X}'\mathbf{V}^{-1}\mathbf{y} \tag{13.47}$$

is the generalized least-squares solution (GLS) of **b and K′b** is the best linear unbiased estimator (BLUE) of **K′b**, given that **K′b** is estimable. BLUE is similar in meaning and properties to BLUP but relates to estimates of linear functions of fixed effects. An estimator of the estimable functions of fixed effects has minimum sampling variance, is unbiased and is based on a linear function of the data (Henderson, 1984).

The predictor, $\mathbf{L}'\mathbf{y}$, is the same as BLP except that $\mathbf{b}$ has been replaced with $\hat{\mathbf{b}}$. Thus only the second moment of $\mathbf{y}$, var($\mathbf{y}$) = $\mathbf{V} = \mathbf{Z}'\mathbf{G}\mathbf{Z} + \mathbf{R}$ is assumed known to compute BLUP. A predictor of $\mathbf{K}'\mathbf{b} + \mathbf{M}'\mathbf{u}$ is $\mathbf{K}'\hat{\mathbf{b}} + \mathbf{M}'\hat{\mathbf{u}}$, which is BLUP only if $\mathbf{G}$ and $\mathbf{R}$ are known and the model is correct (Henderson, 1973). A predictor of $\mathbf{u}$ is obtained by letting $\mathbf{K}' = \mathbf{0}$ and $\mathbf{M}' = \mathbf{I}$, then

$$\mathbf{L}'\mathbf{y} = \hat{\mathbf{u}} = \text{BLUP}(\mathbf{u}) = \mathbf{G}\mathbf{Z}'\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\hat{\mathbf{b}}) \tag{13.48}$$

The true value of $\mathbf{G}$ and $\mathbf{R}$ are never known in practical animal breeding. However, if $\mathbf{G}$ and $\mathbf{R}$ are replaced by estimates obtained by one of the usual variance component estimation methods (see Chapter 9), then use of those estimates yields unbiased estimate of $\mathbf{b}$ and unbiased predictor of $\mathbf{u}$ provided that $\mathbf{y}$ is normally distributed (Kackar and Harville, 1981).

The equation to compute BLUP of $\mathbf{K}'\mathbf{b} + \mathbf{M}'\mathbf{u}$ requires $\mathbf{V}^{-1}$, which is generally of order *n*, the number of elements in $\mathbf{y}$. In most situations in animal breeding *n* is very large and the inversion of $\mathbf{V}$ becomes virtually impossible. However, Henderson (1950) discovered the mixed model equations (MME) as an alternative computational formula for BLUP of $\mathbf{u}$ and GLS of $\mathbf{b}$. However, proof that MME gives GLS of $\mathbf{b}$ was not shown until 1961 with the help of S. R. Searle, and the proof that selection index and the solution to $\mathbf{u}$ in MME were the same was established shortly thereafter (Henderson, 1963).

The MME for $\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{Z}\mathbf{u} + \mathbf{e}$ is:

$$\begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{b}} \\ \hat{\mathbf{u}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{y} \end{bmatrix} \tag{13.49}$$

assuming that $\mathbf{R}$ and $\mathbf{G}$ are non-singular (unique inverse exist) and requires the inverse of $\mathbf{R}$, which is also of order n the same as $\mathbf{V}$. Usually $\mathbf{R}$ is diagonal and can be factorised from both sides of the equation to give:

$$\begin{bmatrix} \mathbf{X}'\mathbf{X} & \mathbf{X}'\mathbf{Z} \\ \mathbf{Z}'\mathbf{X} & \mathbf{Z}'\mathbf{Z} + \mathbf{G}^{-1}\alpha \end{bmatrix} \begin{bmatrix} \hat{\mathbf{b}} \\ \hat{\mathbf{u}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{y} \\ \mathbf{Z}'\mathbf{y} \end{bmatrix} \tag{13.50}$$

with $\alpha = \sigma_e^2 / \sigma_a^2 = (1 - h^2)/h^2$.

Also the inverse of $\mathbf{G}$ is needed which is a matrix of order equal to the number of elements in $\mathbf{u}$. The ability to compute $\mathbf{G}^{-1}$ depends on the model and the definition of $\mathbf{u}$. In animal breeding data $\mathbf{u}$ may represent a vector of animal additive genetic effects. With unrelated animals $\mathbf{G}^{-1} = \mathbf{I}\frac{1}{\sigma_a^2}$, while for related animals $\mathbf{G}^{-1} = \mathbf{A}^{-1}\frac{1}{\sigma_a^2}$, where $\mathbf{A}^{-1}$ is

the inverse of the additive genetic relationship matrix (see section 13.7 for how to construct **A** and $\mathbf{A}^{-1}$).

The order of the MME equals the number of elements in **b** (p) plus the number of elements in **u** (q). Frequently (p + q) is considerably less than n, but could still be large enough to prevent the calculation of the inverse of the left hand side of MME, which is required to obtain **b** and **u** directly. However, the MME can be solved by iteration (e.g. Gauss-Seidel) or by Markov chain Monte Caro (MCMC) methods (e.g. Gibbs sampler; Casella and George, 1992; Sorensen et al., 1994). From the latter, but not the former, the variances of prediction error can be obtained without inverting the left hand side of the MME.

The variances of the estimated fixed effects are obtained as:

$$V(\hat{\mathbf{b}}) = \mathbf{C_{11}} \tag{13.51}$$

,

the variance of the predicted random effects as

$$V(\hat{\mathbf{u}}) = Cov(\hat{\mathbf{u}}, \mathbf{u}') = \mathbf{G} - \mathbf{C_{22}} \tag{13.52}$$

and the prediction error variance of the random effects as:

$$PEV = V(\hat{\mathbf{u}} - \mathbf{u}) = V(\hat{\mathbf{u}}) + V(\mathbf{u}) - Cov(\hat{\mathbf{u}}, \mathbf{u}') - Cov(\mathbf{u}, \hat{\mathbf{u}}') = V(\mathbf{u}) - V(\hat{\mathbf{u}}) = \mathbf{C_{22}} \tag{13.53}$$

where $\mathbf{C_{11}}$ and $\mathbf{C_{22}}$ are the elements of the inversed coefficient matrix pertaining to the fixed and random effects, respectively. Note that $PEV = (1 - r^2_{A\hat{A}})V(\mathbf{u})$ (see section 13.4.7).

13.6 Differences between SIP and BLUP

Selection Index Procedure assumes that the records have the same means or that the means are known and also that the variances are known. In BLUP the means for the fixed effects (BLUE) are obtained simultaneously with the random effects (BLUP) and hence only the variances are assumed known. BLUP is selection index with the means for fixed effects replaced with the GLS (Generalized Least Squares) estimates for these effects.

To illustrate the differences between the selection index and BLUP, first consider the general equations for selection index (Best Linear Predictor):

$$E(A_i | P_i) = \hat{A}_i = \hat{\mu}_A + \frac{\sigma_{AP}}{\sigma^2_P} P_i^{\otimes} \tag{13.54}$$

where $P_i^{\otimes}$ is the observed record for animal *i* adjusted for some fixed effects and the other parameters are defined earlier. For BLUP the comparable equation is:

$$E(A_i|P_i) = \hat{A}_i = \hat{\mu}_A + \hat{g}_i + \frac{\sigma_{AP}}{\sigma_P^2}(P_i - f(\hat{a}, \hat{b},.., \hat{\mu}_i,..) \quad (13.55)$$

where $\hat{\mu}_A + \hat{g}_i$ is BLUE of the genetic level of all animals in the population, $f(\hat{a}, \hat{b},.., \hat{\mu}_i,..)$ is BLUE for the fixed effects (a,b, ..) associated with the observed phenotype and BLUP for the animals related to the animal in question.

Use of BLUP accounts for changes in genetic variance due to accumulated relationships or inbreeding if **A** is complete. Changes in genetic variance from assortative (or disassortative) mating are also accounted for, because with BLUP under an animal model adjustment for the genetic merit of the mate is made. In addition selection is accounted for if the complete pedigree is included.

13.7 Prediction of breeding values with BLUP

To illustrate the use of BLUP, consider the data in Table 13.1 where body weight records are available on 10 fish (i.e. 6 males and 4 females of the same age) and raised in two cages. The fish were the offspring of two sires and four dams. The parents of the sires and dams are all known, while the parents of the grandparents are unknown. Thus, the pedigree file for the example data is shown in Table 13.2.

Table 13.1. Example data

Fish id	Cage	Sex	Sire id	Dam id	Body weight, kg
21	1	M	11	13	6.1
22	1	M	11	14	5.4
23	1	M	12	15	3.6
24	1	F	11	13	5.5
25	1	F	12	15	5.7
26	1	F	12	16	4.9
27	2	M	11	13	7.8
28	2	M	12	15	9.3
29	2	M	12	16	6.6
30	2	F	11	13	6.9

In matrix notation the sire and dam model may be written as:

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Z_s u_s} + \mathbf{Z_d u_d} + \mathbf{e} \quad (13.56)$$

Where $\mathbf{y}$ = is a 10x1 vector of observed body weights
$\mathbf{b}$ = is a 5x1 vector of fixed effects; the overall mean with one level and cage and sex with two levels each;
$\mathbf{u_s}$ = 2 x 1 vector of ½ the sire additive genetic values
$\mathbf{u_d}$ =4 x 1 vector of ½ the dam additive genetic values

e = 10 x 1 vector of residual errors; i,e., the random environmental effects and Mendelian sampling values, **X,** $\mathbf{Z_s}$ and $\mathbf{Z_d}$ are the corresponding design matrices of order 10 x 5, 10 x 2 and 10 x 4, respectively.

The above coding of the two fixed effects cage and sex assumes that the magnitude of the sex effect is the same for both cages. This is not necessarily the truth, in particular when there are large differences in the mean trait value between the cages. This can happen if the fish are reared at two different farms or if the records in the two cages are recorded at a different time of the year. In that case a combined cage by sex effect with four levels (c_1s_1, c_1s_2, c_2s_1, c_2s_2) should be used.

Table 13.2. The pedigree of the example data in Table 13.1.

Fish id	Sire id	Dam id	Generation
1	0	0	0
2	0	0	0
3	0	0	0
4	0	0	0
5	0	0	0
6	0	0	0
7	0	0	0
8	0	0	0
11	1	4	1
12	2	5	1
13	2	5	1
14	3	6	1
15	1	7	1
16	3	8	1
21	11	13	2
22	11	14	2
23	12	15	2
24	11	13	2
25	12	15	2
26	12	16	2
27	11	13	2
28	12	15	2
29	12	16	2
30	11	13	2

The expectation of the observation vector is $\mathbf{E(y) = Xb}$. The expectation and variance and covariance matrix for the random effects are:

$$\mathbf{E}\begin{bmatrix}\mathbf{u_s}\\ \mathbf{u_d}\\ \mathbf{e}\end{bmatrix}=\begin{bmatrix}\mathbf{0}\\ \mathbf{0}\\ \mathbf{0}\end{bmatrix} \text{ and } \mathbf{E}\begin{bmatrix}\mathbf{u_s}\\ \mathbf{u_d}\\ \mathbf{e}\end{bmatrix}=\begin{bmatrix}\mathbf{A_{ss}}\sigma_a^2 & \mathbf{A_{sd}}\sigma_a^2 & 0\\ \mathbf{A_{ds}}\sigma_a^2 & \mathbf{A_{dd}}\sigma_a^2 & 0\\ 0 & 0 & 0\end{bmatrix}, \tag{13.56}$$

where $\sigma_a^2 = 0.25\sigma_A^2$, i.e., one quarter of the additive genetic variance and **A** is the additive genetic relationship matrix among the sires ($\mathbf{A_{ss}}$), among the dams ($\mathbf{A_{dd}}$) and among the sires and dams ($\mathbf{A_{sd}}$ and $\mathbf{A_{ds}}$). If the observed trait is influenced by effects other than additive genetics (e.g., non-additive genetic, environmental effects common to full-sibs and maternal effects), these effects will be confounded with the dam additive genetic value.

The design matrices can be easily constructed as:

$$\mathbf{X}=\begin{bmatrix}1&1&0&1&0\\1&1&0&1&0\\1&1&0&1&0\\1&1&0&0&1\\1&1&0&0&1\\1&0&1&0&1\\1&0&1&1&0\\1&0&1&1&0\\1&0&1&1&0\\1&0&1&0&1\end{bmatrix} \quad \mathbf{Z_s}=\begin{bmatrix}1&0\\1&0\\0&1\\1&0\\0&1\\0&1\\1&0\\0&1\\0&1\\1&0\end{bmatrix} \quad \mathbf{Z_d}=\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\1&0&0&0\\0&0&1&0\\0&0&0&1\\1&0&0&0\\0&0&1&0\\0&0&0&1\\1&0&0&0\end{bmatrix}$$

In **X** the first column pertain to the overall mean, the second and third columns to the two levels of the cage effect, and the fourth and fifth columns to the two levels of the sex effect. The two columns of $\mathbf{Z_s}$ pertain to the two levels of the sire effect while the four columns of $\mathbf{Z_d}$ pertain to the four levels of the dam effect.

The normal equations are:

$$\begin{bmatrix}\mathbf{X'X} & \mathbf{X'Z_s} & \mathbf{X'Z_d}\\ \mathbf{Z_s'X} & \mathbf{Z_s'Z_s} & \mathbf{Z_s'Z_d}\\ \mathbf{Z_d'X} & \mathbf{Z_d'Z_s} & \mathbf{Z_d'Z_d}\end{bmatrix}\begin{bmatrix}\mathbf{b}\\ \hat{\mathbf{u}}_\mathbf{s}\\ \hat{\mathbf{u}}_\mathbf{d}\end{bmatrix}=\begin{bmatrix}\mathbf{X'y}\\ \mathbf{Z_s'y}\\ \mathbf{Z_d'y}\end{bmatrix} \tag{13.58}$$

With the example data the normal equations become:

$$\begin{bmatrix} 6 & 0 & 3 & 3 & 3 & 3 & 2 & 1 & 2 & 1 \\ & 4 & 3 & 1 & 2 & 2 & 2 & 0 & 1 & 1 \\ & & 6 & 0 & 3 & 3 & 2 & 1 & 2 & 1 \\ & & & 4 & 2 & 2 & 2 & 0 & 1 & 1 \\ & & & & 5 & 0 & 4 & 1 & 0 & 0 \\ & & & & & 5 & 0 & 0 & 3 & 2 \\ & & sym & & & & 4 & 0 & 0 & 0 \\ & & & & & & & 1 & 0 & 0 \\ & & & & & & & & 3 & 0 \\ & & & & & & & & & 2 \end{bmatrix} \begin{bmatrix} \hat{c}_1 \\ \hat{c}_2 \\ \hat{s}_1 \\ \hat{s}_2 \\ \hat{u}_{11} \\ \hat{u}_{12} \\ \hat{u}_{13} \\ \hat{u}_{14} \\ \hat{u}_{15} \\ \hat{u}_{16} \end{bmatrix} = \begin{bmatrix} 31.2 \\ 30.6 \\ 38.8 \\ 23.0 \\ 31.7 \\ 30.1 \\ 26.3 \\ 5.4 \\ 18.6 \\ 11.5 \end{bmatrix}$$

Here the row and column for the overall mean have been omitted and the vector for the fixed effects (**b**) is partitioned into its components for cage ($\mathbf{c_1}$ and $\mathbf{c_2}$) and sex ($\mathbf{s_1}$ and $\mathbf{s_2}$).

To obtain BLUP for sires and dams, these equations where all effects are fixed, have to be converted into equations for a mixed model (Henderson's mixed model equations, MME). This is obtain by adding $\mathbf{G}^{-1}$, i.e., the inverse of the additive genetic variance-covariance matrix, to the random effect parts of the above equations, i.e.:

$$\mathbf{Z}'\mathbf{Z} = \begin{bmatrix} \mathbf{Z}_s'\mathbf{Z}_s & \mathbf{Z}_s'\mathbf{Z}_d \\ \mathbf{Z}_d'\mathbf{Z}_s & \mathbf{Z}_d'\mathbf{Z}_d \end{bmatrix} \tag{13.59}$$

If all sires and dams were unrelated we would add the following matrix to $\mathbf{Z}'\mathbf{Z}$:

$$\mathbf{G}^{-1} = \begin{bmatrix} \mathbf{I}k & \mathbf{0} \\ \mathbf{0} & \mathbf{I}k \end{bmatrix}, \text{ where } k = \frac{\sigma_e^2}{\sigma_s^2} = \frac{\sigma_e^2}{\sigma_d^2} = \frac{1 - 1/2\,h^2}{1/4h^2} \text{ when assuming}$$

that both the sire and dam component of variance is equal to one quarter of the additive genetic variance. The magnitude of the sire and dam component of variance could be different.

However, since the sires and dams are related:

$$\mathbf{G}^{-1} = k\mathbf{A}^{-1} = k \begin{bmatrix} \mathbf{A}^{s,s} & \mathbf{A}^{s,d} & \mathbf{A}^{s,ss} & \mathbf{A}^{s,dd} \\ \mathbf{A}^{d,s} & \mathbf{A}^{d,d} & \mathbf{A}^{d,ss} & \mathbf{A}^{d,dd} \\ \mathbf{A}^{ss,s} & \mathbf{A}^{ss,d} & \mathbf{A}^{ss,ss} & \mathbf{A}^{ss,dd} \\ \mathbf{A}^{dd,s} & \mathbf{A}^{dd,d} & \mathbf{A}^{dd,ss} & \mathbf{A}^{dd,dd} \end{bmatrix} \tag{13.60}$$

where $\mathbf{A}^{s,s}, \mathbf{A}^{s,d}$, etc. are the inverse parts of the additive genetic relationship matrix $\mathbf{A}_{s,s}, \mathbf{A}_{s,d}$, etc., respectively and where the super and subscripts s and d refers to sires and dams, respectively and ss and dd to grandsires and granddames, respectively. When the animals id's are ordered as 11, 12, 13, 14, 15, 16, 1, 2, 3, 4, 5, 6, 7, 8, the elements of $\mathbf{A}$ and $\mathbf{A}^{-1}$ are:

$$\mathbf{A} = \begin{bmatrix}
1 & & & & 1/4 & & 1/2 & & & 1/2 & & & & \\
 & 1 & 1/2 & & & & & 1/2 & & & 1/2 & & & \\
 & & 1 & & & & & 1/2 & & & 1/2 & & & \\
 & & & 1 & & & & & 1/2 & & & 1/2 & & \\
 & & & & 1 & 1/4 & & & & & & & 1/2 & \\
 & & & & & 1 & 1/2 & & 1/2 & & & & & 1/2 \\
 & & & & & & 1 & & & & & & & \\
 & & & & & & & 1 & & & & & & \\
 & & & & & & & & 1 & & & & & \\
 & & & sym & & & & & & 1 & & & & \\
 & & & & & & & & & & 1 & & & \\
 & & & & & & & & & & & 1 & & \\
 & & & & & & & & & & & & 1 & \\
 & & & & & & & & & & & & & 1
\end{bmatrix}$$

$$\mathbf{A}^{-1} = \begin{bmatrix}
2 & & & & & & -1 & & & -1 & & & & \\
 & 2 & & & & & & -1 & & & -1 & & & \\
 & & 2 & & & & & -1 & & & -1 & & & \\
 & & & 2 & & & & & -1 & & & -1 & & \\
 & & & & 2 & & -1 & & & & & & -1 & \\
 & & & & & 2 & & & -1 & & & & & -1 \\
 & & & & & & 2 & & & 1/2 & & & 1/2 & \\
 & & & & & & & 2 & & & 1 & & & \\
 & & & & & & & & 2 & & & 1/2 & & 1/2 \\
 & & & sym & & & & & & 1.5 & & & & \\
 & & & & & & & & & & 2 & & & \\
 & & & & & & & & & & & 1.5 & & \\
 & & & & & & & & & & & & 1.5 & \\
 & & & & & & & & & & & & & 1.5
\end{bmatrix}$$

The elements of $\mathbf{A}^{-1}$ can be obtained in at least three ways: (1) first construct $\mathbf{A}$ using the tabular method (Henderson, 1976) and then inverting it, which can be difficult for large number of animals. From this $\mathbf{A}$, the inbreeding coefficient for each animal can be obtained as the diagonal element minus one; (2) construct $\mathbf{A}^{-1}$ directly without taking inbreeding into account using Henderson's rapid method (Henderson, 1976); or (3) constructing $\mathbf{A}^{-1}$ directly taking inbreeding into account (Quaas, 1976; Meuwissen and Lou, 1992). To obtain $\mathbf{G}^{-1}$ is therefore not a problem. Thus, after adding $\mathbf{G}^{-1}$ the MME becomes:

$$\begin{bmatrix} \mathbf{X'X} & \mathbf{X'Z_s} & \mathbf{X'Z_s} & \mathbf{0} & \mathbf{0} \\ \mathbf{Z_s'X} & \mathbf{Z_s'Z_s} + k\mathbf{A}^{ss} & \mathbf{Z_s'Z_d} + k\mathbf{A}^{sd} & k\mathbf{A}^{s,ss} & k\mathbf{A}^{s,dd} \\ \mathbf{Z_d'X} & \mathbf{Z_d'Z_s} + k\mathbf{A}^{ds} & \mathbf{Z_d'Z_d} + k\mathbf{A}^{dd} & k\mathbf{A}^{d,ss} & k\mathbf{A}^{d,dd} \\ \mathbf{0} & k\mathbf{A}^{ss,s} & k\mathbf{A}^{ss,d} & k\mathbf{A}^{ss,ss} & k\mathbf{A}^{ss,dd} \\ \mathbf{0} & k\mathbf{A}^{dd,s} & k\mathbf{A}^{dd,d} & k\mathbf{A}^{dd,ss} & k\mathbf{A}^{dd,dd} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{b}} \\ \hat{\mathbf{u}}_s \\ \hat{\mathbf{u}}_d \\ \hat{\mathbf{u}}_{ss} \\ \hat{\mathbf{u}}_{dd} \end{bmatrix} = \begin{bmatrix} \mathbf{X'y} \\ \mathbf{Z_s'y} \\ \mathbf{Z_d'y} \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix} \quad (13.51)$$

After adding $\mathbf{G}^{-1} = k\mathbf{A}^{-1}$ for k = 14, i.e., h^2=0.25, the diagonal elements of $\mathbf{Z'Z}$ and the elements of $\mathbf{Z'y}$ become:

$\mathbf{Z'Z}$	$\mathbf{Z'y}$
5 + 2(14) = 33	31.7
5 + 2(14) = 33	30.1
4 + 2(14) = 32	26.3
1 + 2(14) = 29	5.4
3 + 2(14) = 31	18.6
2 + 2(14) = 30	11.5
0 + 2(14) = 28	0
0 + 2(14) = 20	0
0 + 2(14) = 28	0
0 + 1.5(14) = 21	0
0 + 2(14) = 28	0
0 + 1.5(14) = 21	0
0 + 1.5(14) = 21	0
0 + 1.5(14) = 21	0

The off diagonal elements of $\mathbf{Z'Z} + k\mathbf{A}^{-1}$ attain the value –1(14) = -14, ½(14) = 7 or 0.

The estimates for the fixed effects (BLUE) and the solutions for the random effects (BLUP) are given in second column of Table 13.3. The third column gives the BLUE and BLUP values assuming a $h^2 = 0.50$ as compared to $h^2 = 0.25$ in the second column. The estimates and solutions in the fourth column are those obtained for $h^2 = 0.25$ assuming that the sires and dams are not related. The following is worthy of note from the results presented in Table 13.3:

13.7.1 BLUP WITH $\mathbf{A}^{-1}$, $h^2 = 0.25$

The following always applies: $\mathbf{1'A}^{-1}\hat{\mu} = 0$. For this example, $\mathbf{1'A}^{-1}$ = (0 0 0 0 0 0 1 1 1 1 1 1 1 1) which says that the sum of the solutions of the base population animals (i.e, the granparents in this example) is zero.

Table 13.3. Estimates of the fixed effects (BLUE for cage and sex) and predictions of the random effects (BLUP for sires, dams and grandparents and full-sib families) for two different values of h^2 and with and without $\mathbf{G}^{-1} = k\mathbf{A}^{-1}$ added to the normal equations.

Effect	With $\mathbf{G}^{-1}$ $h^2 = 0.25$	With $\mathbf{G}^{-1}$ $h^2 = 0.50$	Without $\mathbf{G}^{-1}$ $h^2 = 0.25$	Least squares estimates[1]
Cages				
$\hat{c}_1$	5.2334	5.2130	5.2400	5.1354
$\hat{c}_2$	7.6642	7.6540	7.6704	7.6771
Sex				
$\hat{s}_1$	0.0000	0.0000	0.0000	0.0000
$\hat{s}_2$	-0.1025	-0.0923	-0.1030	-0.0083
Sires				
$\hat{u}_{11}$	0.0509	0.0913	0.0378	0.1792
$\hat{u}_{12}$	-0.0254	-0.0455	-0.0378	-0.1408
Dams				
$\hat{u}_{13}$	0.0119	0.0170	0.0297	-0.0008
$\hat{u}_{14}$	-0.0105	-0.0183	0.0081	0.0992
$\hat{u}_{15}$	0.0467	0.0952	0.0392	0.3692
$\hat{u}_{16}$	-0.0758	-0.1444	-0.0770	-0.5058
Grandparents				
$\hat{u}_1$	0.0390	0.0746	-	-
$\hat{u}_2$	-0.0045	-0.0095	-	-
$\hat{u}_3$	-0.0345	-0.0651	-	-
$\hat{u}_4$	0.0209	0.0360	-	-
$\hat{u}_5$	-0.0045	-0.0095	-	-
$\hat{u}_6$	0.0045	0.0095	-	-
$\hat{u}_7$	0.0181	0.0386	-	-
$\hat{u}_8$	-0.0390	-0.0746	-	-
Error, $\hat{\sigma}_e^2$	1.1471	1.1229	1.1457	1.7725
Full-sib family				
$\hat{u}_{11} + \hat{u}_{13}$	0.0628	0.1083	0.0675	0.1782
$\hat{u}_{11} + \hat{u}_{14}$	0.0404	0.0730	0.0460	0.2782
$\hat{u}_{12} + \hat{u}_{15}$	0.0213	0.0497	0.0014	0.2282
$\hat{u}_{12} + \hat{u}_{16}$	-0.1012	-0.1899	-0.1148	-0.6468

[1] Obtained by setting $k = \dfrac{\sigma_e^2}{\sigma_s^2} = \dfrac{1 - 1/2\,h^2}{1/4h^2}$ close to zero.

13.7.2 BLUP WITH $\mathbf{A^{-1}}$, $h^2 = 0.50$

In absolute value the solutions for all animals are larger than the solutions for $h^2 = 0.25$. Thus, with increasing heritability, the solutions are less regressed towards the population mean. Consequently, the error variance for $h^2 = 0.50$ is smaller than for $h^2 = 0.25$.

13.7.3 BLUP WITHOUT $\mathbf{A^{-1}}$, $h^2 = 0.25$

Here the sires and dams represent the base population animals and therefore since $\mathbf{1'A^{-1}}\hat{\mu} = 0$ always applies, $\mathbf{1'A^{-1}} = (1\ 1\ 1\ 1\ 1\ 1)$ and therefore $\sum_{i=1}^{6} \hat{\mu}_i = 0$.

Notice also that $\hat{\mu}_{11} = \hat{\mu}_{13} + \hat{\mu}_{14}$ and $\hat{\mu}_{12} = \hat{\mu}_{15} + \hat{\mu}_{16}$; hence the dam solutions within sire sum to the solution for the sire.

13.7.4 LEAST SQUARES ESTIMATES

In absolute value the least squares estimates for sires and dams are much larger (unregressed) than the corresponding solutions obtained from the mixed model. The error variance for the fixed model is much larger than for the mixed models. Worth noticing is that for this fixed model the error degree of freedom is 10 – 6 = 4 while it is 10 – 3 = 7 for the mixed model.

13.7.5 BREEDING VALUES FOR THE PROGENY

From the above sire and dam model breeding values for progeny were not obtained. The estimates and solutions from the sire and dam model can be used with the following model to obtain progeny solutions. Let

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Z_s u_s} + \mathbf{Z_d u_d} + \mathbf{Z_m u_m} + \mathbf{e} \tag{13.62}$$

where $\mathbf{u_m}$ is a 10x1 vector of individual Mendelian sampling values with variance $\sigma_w^2 = 0.50\sigma_A^2$, i.e., half of the additive genetic variance, and $\mathbf{Z_m} = \mathbf{I}$, i.e., an identity matrix because there is only one observation per animal. The remaining symbols are as defined earlier.

A prediction of $\mathbf{u_m}$ is:

$$\hat{\mathbf{u}}_\mathbf{m} = (\mathbf{Z_m'Z_m} + \mathbf{Im})^{-1}(\mathbf{y} - \mathbf{X}\hat{\mathbf{b}} - \mathbf{Z_s}\hat{\mathbf{u}}_\mathbf{s} - \mathbf{Z_d}\hat{\mathbf{u}}_\mathbf{d})^{-1} \tag{13.63}$$

where $m = \dfrac{(1 - h^2)}{1/2\,h^2}$ and $\mathbf{Z_m'Z_m} = \mathbf{I}$ (m=6 for $h^2 = 0.25$).

Thus a prediction of the Mendelian sampling value for fish_id = 21 in the example is (h^2 = 0.25 and with $\mathbf{A^{-1}}$ included):

$$\hat{m}_{21} = \frac{1}{1+m}(y - (\hat{c}_1 + \hat{c}_2) - (\hat{u}_{11} + \hat{u}_{13})) = \frac{1}{1+6}(6.1 - (5.2334 + 0) - (0.0509 + 0.0119)) = 0.1148$$

The value $\hat{m}_{21}$ is composed by, (1) adjusting the observed phenotypic value for the fixed effects and (2) expressing this value as a deviation from the mean breeding value of its full-sibs. Therefore, the predicted breeding value of fish_id = 21 is $EBV_{21} = (\hat{u}_{11} + \hat{u}_{13}) + \hat{m}_{21} = 0.0628 + 0.1148 = 0.1776$. The predicted breeding values for the other fish in the example are shown in Table 13.2.

13.7.6 THE EFFECT OF A^{-1}

To illustrate the effect of the additive genetic relationship matrix on the predicted breeding value, first consider the normal equation for sire id =11 after adding $k\mathbf{A}^{-1}$:

$$3\hat{c}_1 + 2\hat{c}_2 + 3\hat{s}_1 + 2\hat{s}_s + (5 + 2k)\hat{u}_{11} + 4\hat{u}_{13} + 1\hat{u}_{14} - k\hat{u}_1 - k\hat{u}_4 = 31.7$$

$$\hat{u}_{11} = \frac{31.7}{5} - \frac{3\hat{c}_1 + 2\hat{c}_2}{5} - \frac{3\hat{s}_1 + 2\hat{s}_2}{5} - \frac{4\hat{u}_{13} + 1\hat{u}_{14}}{5} - \frac{2k}{5}[\hat{u}_{11} - 1/2(\hat{u}_1 + \hat{u}_4)]$$

Thus, ½ of the predicted breeding value for the sire ($\hat{u}_{11}$) is composed of (1) the average of its progeny records adjusted for (2) a weighted average of the estimates of two fixed effects (c_i and s_i) and for (3) a weighted average of the solutions for its mates and for (4) the deviation of $\hat{u}_{11}$ from a 'pedigree index' based on its parents. The 'pedigree index' is regressed by a factor that decreases with increasing heritability and increasing number of observations. Thus, the influence of the 'pedigree index' decreases as the progeny information on the breeding candidate increases.

Similarly the normal equation for dam_id = 13 after adding $k\mathbf{A}^{-1}$ is:

$$2\hat{c}_1 + 2\hat{c}_2 + 2\hat{s}_1 + 2\hat{s}_s + 4\hat{u}_{11} + (4 + 2k)\hat{u}_{13} - k\hat{u}_2 - k\hat{u}_5 = 26.3$$

$$\hat{u}_{13} = \frac{26.3}{4} - \frac{2\hat{c}_1 + 2\hat{c}_2}{4} - \frac{2\hat{s}_1 + 2\hat{s}_2}{4} - \hat{u}_{11} - \frac{2k}{4}[\hat{u}_{13} - 1/2(\hat{u}_2 + \hat{u}_5)]$$

13.7.7 SCALING THE PREDICTED VALUES

The predicted breeding values of individuals are considered as random effects and consequently have an expected mean value of zero. For reporting purposes a different mean can be chosen by adding for example the overall mean of the trait.

When the breeding objective includes more than one trait, the overall predicted breeding values are often standardised to a given mean (e.g. 100) and variance (e.g. 10). This is obtained by substracting from each breeding value the overall mean breeding value, dividing by the standard deviation of the breeding values and multiplying with the new standard deviation. The resulting value is added to the new mean. Thus

$$EBV_i^{\oplus} = 100 + \frac{(EBV_i - \overline{EBV})}{\hat{\sigma}_{EBV}} \sigma_{EBV^{\oplus}} \qquad 13.64)$$

where $EBV_i^{\otimes}$ is the standardised breeding value and σ_{EBV} is the standard deviation of the breeding values on the new sale. Re-scaled breeding values become independent of changes caused by shifts in the economic values and population mean for the traits included in the breeding objective and which may take place from one round of selection to the next.

For fish_id 21 in the example the standardised breeding value is:

$$EBV_{21}^{\oplus} = 100 + \frac{0.1776 - 0.0153}{0.1649} 10 = 109.9$$

The standardised breeding values for the other fish in the example are shown in Table 13.4.

Table 13.4. Predicted breeding values for the sires (Sire-id), dams (Dam_id) and the ten offspring (Fish_id) in Table 13.1 ($h^2 = 0.25$ and $\mathbf{A}^{-1}$ included).

Fish_id	Sire_id	Dam_id	Breeding value				
			Sire (u_s)	Dam (u_d)	m_i	EBV_i	$EBV_i^{\oplus}$
21	11	13	0.0509	0.0119	0.115	0.178	109.9
22	11	14	0.0509	-0.0105	0.018	0.058	102.6
23	12	15	-0.0254	0.0467	-0.236	-0.215	86.0
24	11	13	0.0509	0.0119	0.044	0.107	105.6
25	12	15	-0.0254	0.0467	0.078	0.100	105.1
26	12	16	-0.0254	-0.0758	-0.019	-0.120	91.8
27	11	13	0.0509	0.0119	0.010	0.073	103.5
28	12	15	-0.0254	0.0467	0.231	0.252	114.4
29	12	16	-0.0254	-0.0758	-0.138	-0.239	84.6
30	11	13	0.0509	0.0119	-0.103	-0.041	96.9

13.7.8 EFFECTS COMMON TO FULL-SIBS OTHER THAN ADDITIVE GENETICS

When physical tags (e.g. PIT tags) are used for pedigree identification, full-sib groups are reared in separate units (e.g. in tanks) until the sibs reach a body size suitable for

tagging. This may introduce an environmental effect common to the full-sibs (tank effect). To account for this tank effect and also for other effects than additive genetics (non-additive genetics and maternal effect) that will be confounded with the tank effect when full-sib groups are not reared in replicated tanks, the following model may be applied:

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Z_s u_s} + \mathbf{Z_d u_d} + \mathbf{Z_d t_d} + \mathbf{e} \qquad (13.65)$$

where $\mathbf{t_d}$ = 4 x 1 vector of the tank effect (i.e., the effect common to full-sibs other than additive genetics) and with elements corresponding to each full-sib group and thus of the same order as the number of tanks (dams) and with design matrix equal to that for dams.

With the example data the normal equations become:

$$\left[\begin{array}{cc:cc:cc:cccc:cc:cccc:c}
6 & 0 & 3 & 3 & 3 & 3 & 2 & 1 & 2 & 1 & \mathbf{0} & \mathbf{0} & 2 & 1 & 2 & 1 & \hat{c}_1 \\
 & 4 & 3 & 1 & 2 & 2 & 2 & 0 & 1 & 1 & 0 & 0 & 2 & 0 & 1 & 1 & \hat{c}_2 \\ \hdashline
 & & 6 & 0 & 3 & 3 & 2 & 1 & 2 & 1 & 0 & 0 & 2 & 1 & 2 & 1 & \hat{s}_1 \\
 & & & 4 & 2 & 2 & 2 & 0 & 1 & 1 & 0 & 0 & 2 & 0 & 1 & 1 & \hat{s}_2 \\ \hdashline
 & & & & 5 & & 4 & 1 & 0 & 0 & 0 & 0 & 4 & 1 & 0 & 0 & \hat{u}_{11} \\
 & & & & & 5 & 0 & 0 & 3 & 2 & 0 & 0 & 0 & 0 & 3 & 2 & \hat{u}_{12} \\ \hdashline
 & & & & & & 4 & 0 & 0 & 0 & 0 & 0 & 4 & 0 & 0 & 0 & \hat{u}_{13} \\
 & & & & & & & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & \hat{u}_{14} \\
 & & & & & & & & 3 & 0 & 0 & 0 & 0 & 0 & 3 & 0 & \hat{u}_{15} \\
 & & & & & & & & & 2 & 0 & 0 & 0 & 0 & 0 & 2 & \hat{u}_{16} \\ \hdashline
 & & & & & & & & & & 0 & 0 & 0 & 0 & 0 & 0 & \hat{\mathbf{u}}_{\mathbf{gs}} \\
 & & & & & & & & & & & 0 & 0 & 0 & 0 & 0 & \hat{\mathbf{u}}_{\mathbf{gd}} \\ \hdashline
 & & & & & & & & & & & & 4 & 0 & 0 & 0 & \hat{t}_{13} \\
 & & & & & & & & & & & & & 1 & 0 & 0 & \hat{t}_{14} \\
 & & & & & & & & & & & & & & 3 & 0 & \hat{t}_{15} \\
 & & & & & & & & & & & & & & & 2 & \hat{t}_{16}
\end{array}\right]
=
\left[\begin{array}{c}
31.2 \\ 30.6 \\ \hdashline 38.8 \\ 23.0 \\ \hdashline 31.7 \\ 30.1 \\ \hdashline 26.3 \\ 5.4 \\ 18.6 \\ 11.5 \\ \hdashline 0 \\ 0 \\ \hdashline 26.3 \\ 5.4 \\ 18.6 \\ 11.5
\end{array}\right]$$

The above equation are those before adding $\mathbf{G}^{-1} = k\mathbf{A}^{-1}$ to the random animal part (parents and grandparents) and $k_2 = \dfrac{\sigma_e^2}{\sigma_t^2} = \dfrac{1 - 1/2\, h^2 - t^2}{t^2}$ to the random tank effect part where σ_t^2 is the variance components for the tank effect and $t^2 = \dfrac{\sigma_t^2}{\sigma_s^2 + \sigma_d^2 + \sigma_t^2 + \sigma_e^2}$.

To illustrate the different elements involved in the estimates of the sire, dam and tank effects, consider first the normal equation for sire_id =11 after adding $\mathbf{G}^{-1} = k\mathbf{A}^{-1}$ to the random animal part (parents and grandparents):

$$3\hat{c}_1 + 2\hat{c}_2 + 3\hat{s}_1 + 2\hat{s}_s + (5+2k)\hat{u}_{11} + 4\hat{u}_{13} + 1\hat{u}_{14} - k\hat{u}_1 - k\hat{u}_4 + 4t_{13} + 1t_{14} = 31.7$$

$$\hat{u}_{11} = \frac{31.7}{5} - \frac{3\hat{c}_1 + 2\hat{c}_2}{5} - \frac{3\hat{s}_1 + 2\hat{s}_2}{5} - \frac{4\hat{u}_{13} + 1\hat{u}_{14}}{5} - \frac{2k}{5}[\hat{u}_{11} - 1/2(\hat{u}_1 + \hat{u}_4)] - \frac{4\hat{t}_{13} + 1\hat{t}_{14}}{5}$$

Similarly the normal equation for dam_id = 13:

$$2\hat{c}_1 + 2\hat{c}_2 + 2\hat{s}_1 + 2\hat{s}_s + 4\hat{u}_{11} + (4+2k)\hat{u}_{13} - k\hat{u}_2 - k\hat{u}_5 + 4t_{13} = 26.3$$

$$\hat{u}_{13} = \frac{26.3}{4} - \frac{2\hat{c}_1 + 2\hat{c}_2}{4} - \frac{2\hat{s}_1 + 2\hat{s}_2}{4} - \hat{u}_{11} - \frac{2k}{4}[\hat{u}_{13} - 1/2(\hat{u}_2 + \hat{u}_5)] - \hat{t}_{13}$$

Thus, ½ of the predicted breeding value for the sire ($\hat{u}_{11}$) and the dam ($\hat{u}_{13}$), are composed of the same elements as without tank effect in the model. However, when a tank effect is included, ½ the predicted breeding value for the sire is reduced by an amount equal to the weighted average of the tank effects in which his offspring were reared, while ½ the predicted breeding value for the dam is reduced by an amount equal to the tank effects in which her offspring were reared.

The normal equation for the tank effect t_{13}; i.e., the effect corresponding to the offspring (full-sibs) of sire_id =11 and dam_id =13, after adding k_2 to the diagonal elements of $\mathbf{Z_t}$ is:

$$2\hat{c}_1 + 2\hat{c}_2 + 2\hat{s}_1 + 2\hat{s}_s + 4\hat{u}_{11} + 4\hat{u}_{13} + (4+k_2)\hat{t}_{13} = 26.3$$

$$\hat{t}_{13} = \frac{1}{1 + \frac{k_2}{4}}\left[\frac{26.3}{4} - \frac{2\hat{c}_1 + 2\hat{c}_2}{4} - \frac{2\hat{s}_1 + 2\hat{s}_2}{4} - (\hat{u}_{11} + \hat{u}_{13})\right]$$

Thus, the solution for the tank effect ($\hat{t}_{13}$) is composed of (1) the average of the observations (2) adjusted for a weighted average of the estimates of two fixed effects (c_i and s_i) and (3) adjusted for the sum of the solutions for the sire and the dam of the full-sibs reared in the tank. The resulting quantity is regressed by a factor that decreases with increasing tank effect and with decreasing number of full-sibs. Thus, for a large tank effect (i.e., when k_2 is small), the difference between the observed tank mean adjusted for the fixed effects and the sire and dam solutions will to a larger extent be regressed towards zero as compared to a case with lower tank effect (i.e., when k_2 is large).

In the case of no tank effect ($\sigma_t^2 = 0$), the regression factor for the tank effect becomes unity and the observed tank mean adjusted for the fixed effects becomes equal to the sum of the sire and dam solutions and consequently $\hat{t}_i = 0$.

For the example data, the estimates for the fixed effects (BLUE) and the solutions for the random effects (BLUP) when tank effect is included in the model are shown in Table 13.5.

Table 13.5. Estimates of the fixed effects (BLUE for cage and sex) and predictions of the random effects (BLUP for sires, dams, grandparents, full-sib families and tank effect) for $h^2 = 0.25$, $t^2=0.10$ and with and without $\mathbf{G}^{-1} = k\mathbf{A}^{-1}$ added to the normal equations.

Effects	With $\mathbf{G}^{-1}$ $t^2 = 0.00$	With $\mathbf{G}^{-1}$ $t^2 = 0.10$	Without $\mathbf{G}^{-1}$ $t^2 = 0.10$
Cages			
$\hat{c}_1$	5.2334	5.2139	5.2192
$\hat{c}_2$	7.6642	7.6570	7.6624
Sex			
$\hat{s}_1$	0.0000	0.0000	0.0000
$\hat{s}_2$	-0.1025	-0.0903	-0.0901
Grandparents			
$\hat{u}_1$	0.0390	0.0352	-
$\hat{u}_2$	-0.0045	-0.0052	-
$\hat{u}_3$	-0.0345	-0.0299	-
$\hat{u}_4$	0.0209	0.0191	-
$\hat{u}_5$	-0.0045	-0.0052	-
$\hat{u}_6$	0.0045	0.0052	-
$\hat{u}_7$	0.0181	0.0161	-
$\hat{u}_8$	-0.0390	-0.0352	-
Sires			
$\hat{u}_{11}$	0.0509	0.0462	0.0350
$\hat{u}_{12}$	-0.0254	-0.0243	-0.0350
Dams			
$\hat{u}_{13}$	0.0119	0.0086	0.0253
$\hat{u}_{14}$	-0.0105	-0.0071	0.0097
$\hat{u}_{15}$	0.0467	0.0417	0.0344
$\hat{u}_{16}$	-0.0758	-0.0677	-0.0694

Tank effect			
$\hat{t}_{13}$	-	0.0442	0.0405
$\hat{t}_{14}$	-	0.0168	0.0155
$\hat{t}_{15}$	-	0.0515	0.0550
$\hat{t}_{16}$	-	-0.1125	-0.1110
Full-sib family			
$\hat{u}_{11} + \hat{u}_{13}$	0.0628	0.0548	0.0603
$\hat{u}_{11} + \hat{u}_{14}$	0.0404	0.0391	0.0447
$\hat{u}_{12} + \hat{u}_{15}$	0.0213	0.0174	-0.0006
$\hat{u}_{12} + \hat{u}_{16}$	-0.1012	-0.0920	-0.1044
Error, $\hat{\sigma}_e^2$	1.1471	1.2000	1.1186

14. GENOTYPE – ENVIRONMENT INTERACTION

TRYGVE GJEDREM

14.1 Introduction

Aquaculture takes place under very different environmental conditions. There may be differences in water temperature, salinity, technology or feed for example. No species of fish or shellfish are adapted to all kinds of environments. Outside the environment to which it is adapted, a particular species may not thrive or may not be able to survive at all. For example, salmon cannot compete with tilapia at high water temperatures while tilapia are not able to grow and survive in cold waters. We therefore talk about warm water species, coldwater species and species suitable for temperate climate. Similarly some species are adapted to freshwater, brackish water or salt water.

Genotype-environment interaction may partly be ascribed to sensitivity of animals to variation in environmental conditions. Departures in environmental condition from those to which a species is adapted can cause the animal to become stressed. The main factors causing stress-situations for the animals are temperature, salinity, handling and water quality both in running and stagnant waters. Type and availability of feed may also cause interaction. We want robust animals, which will tolerate fluctuations in environmental conditions. In order to avoid developing strains sensitive to different environmental conditions, breeding values should be estimated on data from animals tested under varying farming conditions representative to the industry.

The simple model where phenotype equals genotype plus environment ($P = G + E$) will only be a rough approximation. The model can be extended to include genotype-environment interaction ($P = G + E + COV._{GE}$). Methods for estimation of the genotype - environment interaction component is shown in chapter 9.

14.2 Causes of interaction

The biological causes of interactions between genotype and environment are many. In short, some genotypes fit exceptionally well in certain environmental niches while poorly into others. This is illustrated in Figure 14.1. The Chinese strain performs better than the other strains in the environments with low growth while it cannot compete with the European and the crossbred strain in the environments with high growth rate. In a statistical sense interactions turn up when the ranking of genotypes are different in different environments.

In some data significant genotype-environment interaction may be shown by differences in scale in different environments, such that genotypic differences may be greater in one environment than in the other. This type of interaction can be removed by suitable transformation of data.

T. Gjedrem (ed.), Selection and Breeding Programs in Aquaculture, 233–242.
© 2005 *Springer. Printed in the Netherlands.*

A more serious form of interaction is when the best genotype in one environment is not the best in another environment. This causes problems for estimation of breeding values and will reduce the efficiency of selection since one can not be sure that genetic improvement made in a high level of feed and management will be transferred to lower level of environmental conditions.

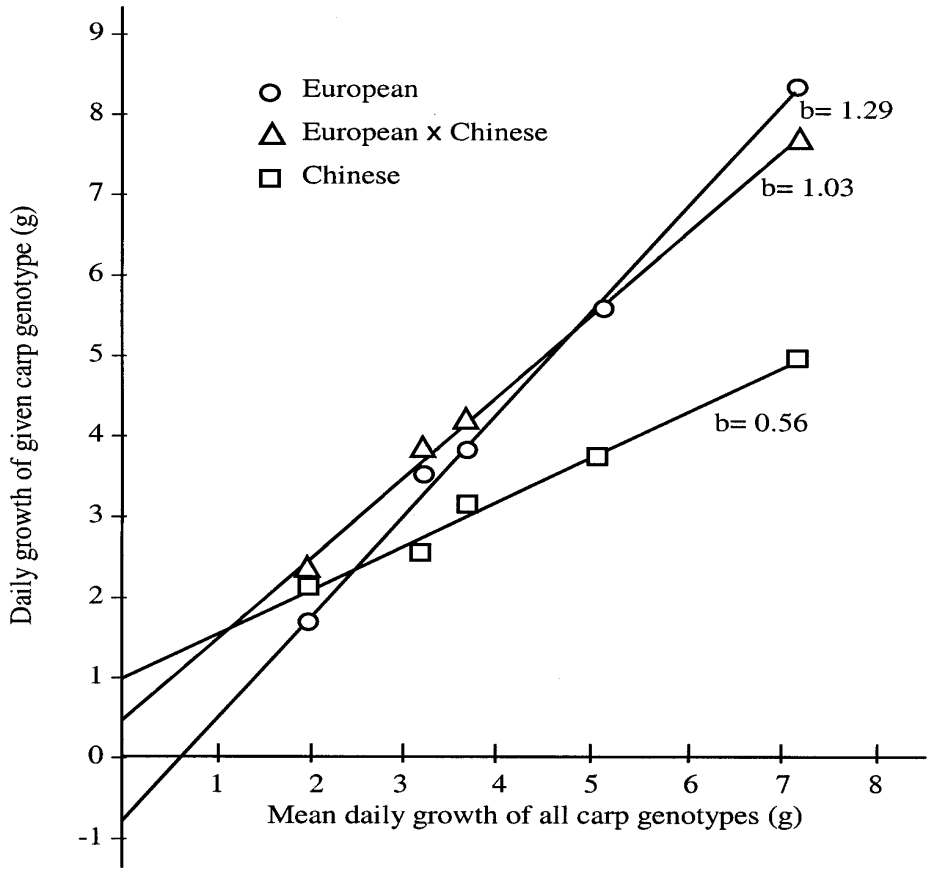


Figure 14.1. Response curves obtained by growing three different genotypes (European, Chinese x European and Chinese) of common carp in different environments. Modified from Wohlfarth et al. (1983) by permission of Elsevier.

According to Pirchner (1985) genotype-environment interactions are important in situations where environment can not be controlled, such as in plant production, or in animal production under range conditions. In contrast, under farming conditions the environmental factors can be partly standardised and then the genotype-environment interactions are usually of less importance.

14.3 Estimation of interaction

When several genotypes are reared in different environments, an analysis of variance in two way classification of genotypes and environments will yield estimates of the variance between genotypes, the variance between the environments, and the variance attributable to interaction of genotype-environment, see Chapter 5. A significant interaction need not be a good assessment of the importance of the interaction. The importance of an interaction can be assessed from it's proportion of the total phenotypic variance, see table 14.1. If the variation due to genotype-environment interaction account for only a small part of the total phenotypic variation, it will cause minor problems in breeding value estimation.

According to Falconer and Mackay (1996) the logical measure for genotype-environment interaction is the genetic correlation between performances of genotypes in two environments. A high genetic correlation indicates interactions i.e the same gene affects the trait in both environments. If genetic effects are present in one environment and only partly repeated in another, the genetic correlation will be low.

14.4 Solving the problem of interaction

The efficiency of a breeding program depends to some extent on the magnitude of genotype-environment interaction. Therefore it is of importance to use some strategies to minimise the interaction. Genotype-environment interaction occurs particularly when the variation in genotypes as well as in environments, is large. When testing families or genetic groups one should avoid large environmental variation. On the other hand one should select environmental conditions for testing families or genetic groups, which are representative for the industry. Already in 1946/47 Haldane and Hammond stated that genotypes would be expressed fully only under a high level of management and recommend that testing of animals for selection should take place in such environments.

If the interaction accounts for a large portion of the phenotypic variation or if the correlation between estimated breeding values in different environments is low, the genetic improvement may become small and even negligible. A possible solution to such situation is to develop a strain for each environmental condition to adapt the animals to their demanded environments.

14.5 Estimates of genotype-environment interaction in aquaculture species

14.5.1 CARP

Twelve different genetic groups of common carp, four strains and their crosses, were reared in five very different pond environments in Israel during 1960 and 1970. Moav et al. (1975) studied genotype-environment interaction for growth rate from these experiments. Growth rate in the best environment was more than twice as high as in the poorest environment. The presence of genotype-environment interaction was evident

from the difference in ranking of the tested groups in the different environments. Especially striking was the difference in ranking of the extreme genotypes between the extreme environments.

In another study, Wohlfarth et al. (1983) compared growth rate of three genotypes of carp, consisting of Chinese and European breed and an inter-racial Chinese x European crossbred. The differences in slopes and intercepts of the response curves of each genotype reared in five different environments demonstrate genotype–environment interaction, Figure 14.1. This demonstrates that, for the trait growth rate, environmental effects greatly influence the relationship between the two genotypes. In a poor environment the Chinese is dominant over the European, in a good environment the European is dominant over the Chinese, while in intermediate environments the inter-racial crossbred was overdominant.

Moav (1976) discussed different causes of genotype-environment interactions. He found that the genotype-season interaction appear to be of considerable magnitude in common carp and conclude that the measure of genetic variability based upon a single season are unreliable. Further he concluded that genotype-husbandry system interaction is of importance for culture systems using stagnant waters and of less importance in running water. He also pointed out that there might be considerable genotype-competition interaction in polyculture.

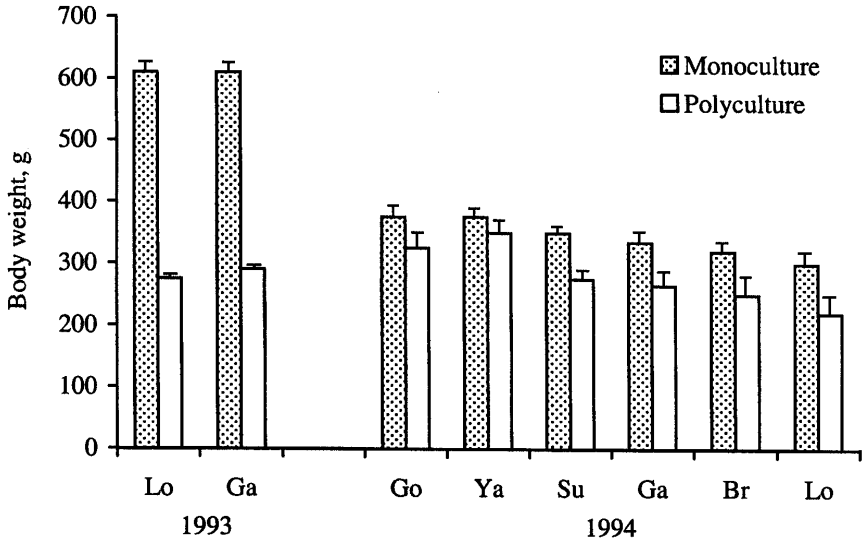


Figure 14.2 Harvest body weight for the stocks in monoculture and polyculture. Ranking based on body weight in monoculture within each year-class. Reproduced from Reddy et al. (2002) by permission of Elsevier.

Genotype-production systems interaction was studied in rohu carp at CIFA (Central Institute of Freshwater Aquaculture), India by Reddy et al. (2001). For the 1993 year-class the interaction in body weight between two strains of rohu carp in mono and

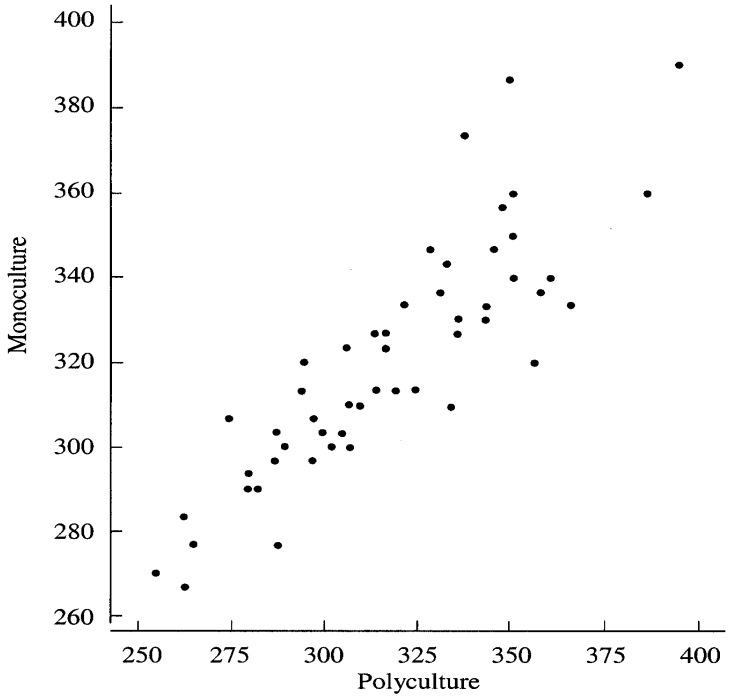


Figure 14.3. Plot of mean breeding values for full-sibs for final body weight (gram) in polyculture and monoculture. Reproduced from Gjerde and Reddy (1996) by permission of Central Institute of Freshwater Aquaculture, India.

polyculture was highly significant but accounted for only 0.1 % of the total variation, Figure 14.2. The interaction in body weight between six strains of rohu in mono and polyculture for body weight was also highly significant in the 1994 year-class, but accounted for a minor portion of the total variation (Fig. 14.2). Similar results were obtained for survival between tagging and harvest.

For the 1994 year-class 57 full-sib groups of rohu were tested in both mono and polyculture. The ranking of the full-sib groups was very similar in these two environments and the genetic correlation between the breeding values in these two environments was high ($r_G = 0.87$) (Gjerde and Reddy, 1996) indicating no significant genotype-environment interaction, Fig. 14.3. This result is contrary to Moav (1976). These results strongly indicate that the growth of rohu may be genetically improved through selection because of high heritability ($h^2 = 0.52$) and based on the test results obtained in either monoculture, polyculture or in both, and that genetically improved seed will show its improved growth potential when stocked in either production system.

14.5.2 SALMONIDES

Gunnes and Gjedrem (1978) studied genotype-environment interaction in Atlantic salmon. Broodstock were sampled from 37 rivers/strains for three consecutive years. Samples of tagged smolts from each strain were distributed to five salmon farms along the Norwegian coast where they were reared in sea cages until reaching market size of 5.45 kg after two years. Total data included 36187 tagged fish. The environmental conditions between the farms were very different. The southern farm was located close to the 61 parallel and the northern close to the 67 parallel with differences in water temperature as well as in day length. In each of the three year-classes the growth rate was only half as good in the farm producing the lowest growth as in the farm producing the highest growth rate. Analysis of variance for body weight and length yielded highly significant strain-farm interaction for each year-class. The component of variance of interaction as a percentage of the total variance explained only a small part, between 1.3 and 3.1 %, of the total variance, as shown in Example below. The ranking of strains at the different farms was very similar. It was therefore concluded that genotype-environment interaction could be neglected in a breeding program for Atlantic salmon and that only one breeding population should be developed for the country.

14.5.2.1 Example

The ANOVA of data from the study by Gunnes and Gjedrem (1978) are shown in Table 14.1. Farm, strain and interaction farm-strain were all highly significant in each of the three year classes. However, the interaction farm-strain accounted for only a small part of the total variance, 3.1%, 2.1% and 1.3% for the years 1972,1973, and 1974, respectively. It is concluded that the genotype-environment interaction is so small that it can be neglected when planning a breeding program for Atlantic salmon.

Table 14.1 ANOVA of body weight in different strains of Atlantic salmon reared in 5 different farms. Giving DF, % of sums of squares for each source of variation, and means of squares (MS). Reproduced from Gunnes and Gjedrem (1978) by permission of Elsevier.

	1972			1973			1974		
Sourse	DF	%	MS	DF	%	MS	DF	%	MS
Farm	4	o,2	10,9**	4	0,1	9,1**	4	1,1	162,8**
Strain	16	6,7	101,6**	22	6,3	75,1**	8	5,2	380,3**
Farm-strain	36	3,1	25,4**	88	2,1	6,3**	21	1,3	37,4**
Residual	6842	89,3	3,1	10449	89,4	2,3	23023	92,0	2,4

Wild et al. (1994), investigated the extent of genotype-environment interaction for age at sexual maturation in salmon. The data consisted of 24 sires and 119 dam progeny groups reared in 6 different net cages at different sites in Norway. The sire–cage and dam-cage interaction accounted for a large part of the total variation. It was concluded that genotype–environment interaction plays an important role for the trait early sexual maturity.

Data of rainbow trout from four year-classes were studied by Gunnes and Gjedrem (1981). The broodstock originally came from different Norwegian fish farms and was treated as one population. Number of sires/progeny groups was 94 with a yearly variation from 18 to 36 genetic groups. Samples of one-year-old fingerlings from each full-sib group were tagged and distributed to five different farms along the coast. A total of 23086 fish reached 2.6 kg after reared in sea cages for 18 months. The differences between

farms were large with about 1 kg difference between the farms with the lowest and the highest body weight of fish. The interaction sire-farm was highly significant but explained only between 1.1 and 5.5 % of the total phenotypic variance for weight and between 0.7 and 4.5 % of the total phenotypic variance for length. Also for rainbow trout it was concluded that genotype-environment interaction could be neglected.

An experiment was carried out in Norway and Sweden to investigate genotype-environment interaction in farmed rainbow trout. In the experiment 35 males and 131 females were mated hierarchically (Sylven et al., 1991). Individual slaughter weights were recorded on 26663 offspring which were reared in eight representative cage culture systems in Norway and Sweden for approximately 1½ years following the first year in fresh water. The slaughter weight was treated as different traits: 1 in fresh water in Sweden; 2 in brackish water in Sweden; and 3 in salt water in Norway. Variance and covariance components for sires, dams within sires and random errors were estimated. Estimates of genetic correlations were 0.86 for traits 2 and 3, 0.72 for traits 1 and 3, and 0.58 for traits 1 and 2. At least for the latter case some genotype-environment interaction must be assumed.

Ayles and Baker (1983) found significant strain-environment (different prairie lakes) in harvest weight of rainbow trout, which could partially be explained by scale effects. This is contrary to Klupp et al. (1978) who concluded that genotype-environment interaction was important for rainbow trout and that scale effects were not present.

14.5.3 TILAPIA

Eight strains of Nile tilapia were tested for two generations in eleven different environments. The test environments were chosen to cover a wide range of Philippine tilapia farming systems, from simple ponds, to more intensive systems, rice-fish systems and cages. During the first generation, trials using individually tagged fingerlings (total 7652) were communally reared in all test environments for about 90 days. In the second generation trials a total of 3420 fingerlings of pure strains were communally reared in eight test environments (Eknath et al., 1993). The strain-environment interaction was significant but accounted for only 0.3 % of the total phenotypic variation in both generations (Eknath et al., 1993). The results from the first generation are illustrated in Figure 14.4. In some of the environments with a very low growth rate, there were some differences in rank-order for some of the strains. From Figure 14.4 it is obvious that the genetic difference between genotypes increases as the environmental conditions measured as growth rate increases.

Genotype-environment interaction was, however, of minor importance in a within family selection experiment for growth rate with Nile tilapia (Bolivar, 1999).

14.5.4 CATFISH

In catfish, Dunham et al. (1990) report that the hybrid between blue catfish and channel catfish grew less than the channel catfish at low densities during the first year of culture in ponds, but that the hybrids grew faster during the second year of culture in ponds.

Genotype-environment interaction appeared when the density was increased or if the culture unit was changed to cages. Similarly, some catfish species and their hybrids performed differently in aquaria than in ponds. A notable example was the blue catfish, which competes with the channel catfish in ponds, lakes and rivers, but performs poorly in aquaria (Sneed, 1971).

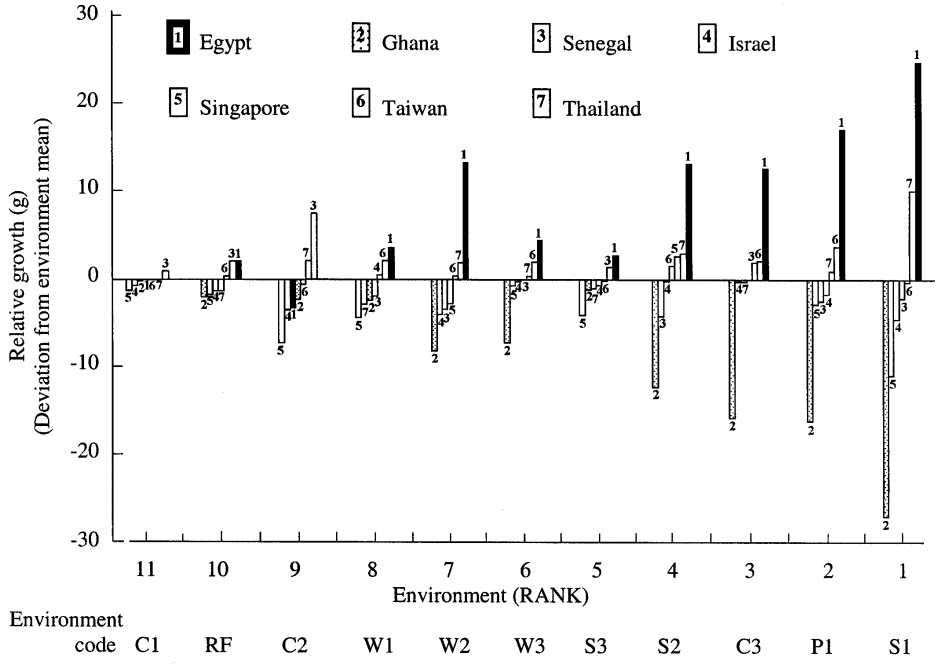


Figure 14.4. Relative growth performance of seven strains of Nile tilapia in different test environments. Relative growth performance has been calculated as a deviation of LSM of final body weights of each strain from the respective environments mean. Reproduced from Eknath, et al. (1993) by permission of Elsevier.

14.5.5 SHRIMP

Oceanic Institute, Hawaii and AKVAFORSK carried out a breeding experiment with *P. vannamei.* Five batches of a total of 294 full-sib groups and close to 18000 tagged progenies were tested in 3 to 4 different farms for growth rate. The family-farm interaction was highly significant for all five batches, but accounted for only 0.5 %, 8 %, 0.1 %, 0.1 % and 0.2 % of the total phenotypic variation for the five batches, respectively. Because genotype-environment interaction was found to be low it is expected that selection based on results from one environment would lead to improved growth in another environment (Fjalestad et al., 1997).

In another experiment with *Penaeus vannamei,* genotype-environment interaction has been studied by CENAIM (Centro Nacional de Acuicultura e Investigaciones) and AKVAFORSK in Columbia. 52 full-sib families were tested for growth rate and pond

survival in two different farms (Suarez, et al., 1999). A strong positive correlation between family growth performance recorded at two farms ($r_G = 0.87$) strongly indicates that the level of genotype-environment interaction for growth is low, Figure 14.5. For pond survival the genetic correlation between family means when reared in two different environments was 0.60 indicating the existence of some genotype-environment interaction.

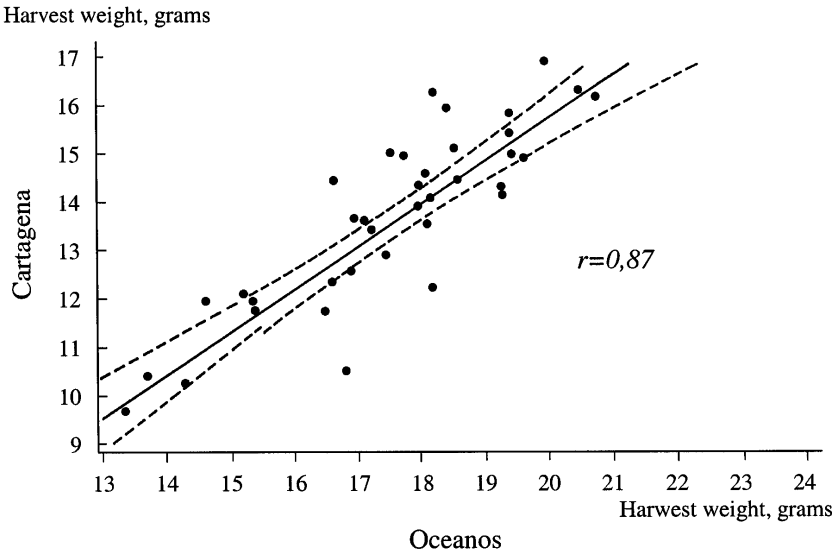


Figure 14.5. Mean family harvest weight of *P. vannamei* at Cartagena and Oceanos. Reproduced from Suarez et al. (1999) by permission of European Aquaculture Society.

14.5.6 MOLLUSCS

For Pacific oyster Langdon et al. (2000) report a strong genotype–environment interaction in body weight although some families performed well across a range of culture environments. Newkirk (1978) compared growth rate of 4 different populations of *Crassostrea virginica* and their crosses in different salinities and found significant genotype–environment interaction.

In scallop Ibarra et al. (1999) did not find significant interaction in growth rate between genotype and tray position.

Jonasson et al. (1999) studied genotype-environment variation in abalone. Progenies of 100 families were reared in two different farms for 14 months. Genotype–environment interaction was negligible for survival and body size.

14.6 Conclusion

These examples of genotype-environment interaction have shown that there is considerable difference in the magnitude of interaction from one study to the next, from one species to another and on the extent of the environmental differences. It is therefore not possible to draw general conclusions about the importance of genotype-environment interaction. Therefore for each species, preferably before the breeding program is established, some estimate of the genotype - environment interaction for the traits of economic importance, and for the environments in which the species will be farmed, should be carried out. Usually the genotype–environment interactions increase with increasing genetic distance and increasing environmental differences. As industry increases productivity, and environments become more standardized across farms, it is likely that the magnitude of interaction will be of less importance.

In general breeders should be tested under environmental conditions, which are representative to the industry. However, one should take into account that some results indicate that testing of breeding values should take place under favourable environmental conditions because of larger genetic variation compared with low environmental level, Figure 14.4. This is in agreement with statements made by Haldane (1946) and Hammond (1947).

Significant interaction causing different ranking of animals must be taken seriously when designing breeding plans. When genotype-environment interaction is considerable, a breeding population should be developed for each type of environment. When interaction results in scale effects it causes small problems in a breeding program because the data may be transformed and the ranking of animals will not be changed.

It has been documented that genotype-environment interaction generally only accounts for a small proportion of the total phenotypic variation although strains with large genetic distance were tested under very different environmental conditions for growth rate and survival in species like salmon, rainbow trout, tilapia, rohu carp and P*enaeus vannamei*.

15. MEASURING GENETIC CHANGE

MORTEN RYE AND TRYGVE GJEDREM

15.1 Introduction

Establishing procedures for monitoring and quantifying genetic changes in a breeding program is not required in order to obtain response to selection, but is essential for several other reasons. Foremost, effective monitoring of the genetic changes that result from the selection process serve as an invaluable tool for assessment of the efficiency of the breeding work, and enables the breeder to make proper adjustments if the realized gains do not meet the theoretical expectations for the current scheme. Valid documentation of genetic progress for key traits of economic importance is also increasingly important for marketing purposes, as providers of eggs, fry or fingerlings often face strong competition in their main markets as aquaculture industries mature.

Implementation of effective protocols for measuring genetic change in breeding programs is not as straightforward as it may intuitively seem. The complexity arises from the fact that changes in the phenotype measured over time are the sum effect of all genetic and environmental changes that occur during the period of interest, and are also affected by possible genotype by environment interaction effects. In this context environmental changes comprise any change of non-genetic nature, and thus are inevitable in any applied breeding program and even in most well controlled selection experiments. Thus, the task of accurately quantifying how much of the change in the animals' phenotypic expression that is resulting solely from the selection process often becomes very complex. Moreover, even with valid response estimates at hand, interpretation of short term responses are further complicated by the fact that initial responses in small populations may fluctuate substantially due to random genetic drift, sampling error, variable selection differentials as well as variable environmental effects (Falconer and MacKay, 1996). Consequently, accurate and reliable estimates of selection response can only be drawn after several cycles of selection.

Estimation of genetic gains in selection programs therefore requires carefully designed protocols in order to provide valid results. An excellent account of alternative approaches with emphasis on application to aquaculture species is given by Gall et al. (1993). In the following, four alternative methods are discussed: the use of an unselected control line, use of divergent selection, repeated matings and genetic trend analysis. A method to quantify the relative selection response originating from the male and female side under a phenotypic selection scheme is also briefly described.

15.2 Control population

Maintaining an unselected, randomly reproduced control line has traditionally been a widely used method for monitoring genetic changes in selection programs. The rationale behind this approach is that any change in environmental conditions is

T. Gjedrem (ed.), Selection and Breeding Programs in Aquaculture, 243–250.
© 2005 *Springer. Printed in the Netherlands.*

expected to equally affect the control and the selected line. The selection response is estimated as the difference in mean performance between the two lines. In order to minimize potential bias due to genotype by environment interaction effects, the control line should ideally be derived from the same genetic base as the selected line. As selection proceeds, however, genotype by environment interactions may still become significant as selection response leads to large differences of the genetic level in the two lines.

Assuming that the control line and the selected line are reared under identical conditions, that the control line is genetically stable and unchanged as compared to the base population for the selected line and no genotype by environment interaction, the difference between the phenotypic mean of the selected and control lines provides an unbiased estimate of the genetic change in the selected line:

$$\Delta G = P_{selected} - P_{control} = (G_{selected} + E_{selected}) - (G_{control} + E_{control}) = (G_{selected} - G_{control}) \quad (15.1)$$

as $E_{control}$ is assumed to be equal to $E_{selected}$. P, G and E refer to phenotypic, genotypic and environmental values, respectively.

Whereas proper management may fulfill the assumption regarding equal rearing condition for the two lines, the genetic stability of the unselected control lines is often questionable. Under most circumstances the effective population size of the control line as well as the selected line is small and thus subject to random drift caused by random change in gene frequencies. Genetic drift affects both the control and selected line, thus limiting the validity of comparisons to a few generations, as discussed by Gall et al. (1993).

Furthermore, the relevance of the application of unselected control lines in breeding programs for aquaculture species may also be questioned on basis of the fact that most fish and shellfish species have a short history of domestication. Even species that have been reproduced in captivity for a long time are still very similar to its wild counterparts, partly because wild broodstock often is regularly used to “refresh” stocks used in culture. The majority of species used in aquaculture are therefore likely to be still under the strong influence of natural selection as they slowly adapt to aquaculture rearing environments which in most cases differ dramatically from the species’ natural environments. Due to the strong forces of natural selection acting in early phases of domestication, the maintenance of a randomly mating and genetically stable control line is likely to be even more difficult for aquaculture species than for livestock species.

Thus, from a theoretical point of view, the use of an unselected control line may not always provide a good means for the estimation of genetic changes resulting from selection. Furthermore, the establishment and maintenance of an unselected control line also carries substantial operational costs and usually necessitates reduced effective population size in the selected line, as the total resources available to the breeding program usually are fixed. Reduced effective population size leads to increased rate of inbreeding at a given selection intensity, or alternatively reduced selection response at a

predefined rate of inbreeding in the selected line. This may represent a substantial indirect cost for the control population method.

15.3 Divergent selection

The relative accuracy of selection response estimates measured by the use of a control line may be increased if a divergent control line is used, i.e. if the control line is selected in the opposite direction as compared with the selected line (Falconer and Mackay, 1996). For this reason, divergent selection has been commonly used to measure response to selection in laboratory experiments. Given that the animals of the two lines are kept under similar environmental conditions and that the selection response is equal in both directions, the genetic change can be estimated as half of the observed difference in performance between the lines.

In most cases when two-way selection is carried out, however, the observed responses are not similar in the upward- and downward-selected lines. This is not surprising, as asymmetry in responses may be caused by number of genetic as well as non-genetic factors (Falconer and MacKay, 1996). Thus, although divergent selection may well be used to demonstrate selection response *per se*, it may be more difficult to accurately quantify genetic changes by this method as compared to the traditional approach involving an unselected control line.

Since divergent selection involves selection in a negative or unfavorable direction in the control line, this method inevitably produces poorly performing control animals for major traits for which selection response is to be estimated (e.g. growth rate). Considering costs, this may represent a substantial disadvantage as compared to use of an unselected control line, which is therefore often preferred for applied programs. Moreover estimates of genetic gain obtained from divergent selection schemes may be severely biased if the trait under selection is affected by genotype by environment interaction as the lines differentiate genetically.

15.4 Repeated matings

Comparison of contemporary animals belonging to different generations provides a means of estimating genetic changes without use of traditional control populations maintained over several generations. In fish and shellfish species representing multiple spawners, repeated matings in successive generation may easily be obtained. For species spawning only once, repeat matings may be facilitated by cryopreservation and storage of frozen sperm.

A simple approach is to compare offspring of males belonging to different generations mated to a random sample of females. Under the assumption that the offspring from both groups of males are reared under the same environmental conditions and that the male side contributes half of the gene pool, the genetic change between the two

generations compared can be measured directly as the difference between mean phenotypic performance of the offspring produced with “new” and “old” sires.

$$P_{new} = \tfrac{1}{2}(G_{newSIRE} + G_{newDAM}) + E_{new}$$

$$\begin{aligned} P_{new} - P_{old} &= [\tfrac{1}{2}(G_{newSIRE} - G_{newDAM}) + E_{new}] - [\tfrac{1}{2}(G_{oldSIRE} - G_{newDAM}) + E_{old}] \\ &= \tfrac{1}{2}(G_{newSIRE} - G_{oldSIRE}), \text{ given that } E_{old} = E_{new} \\ &= \tfrac{1}{2}\Delta G \end{aligned}$$

Thus:
$$\Delta G = 2(P_{new} - P_{old}) \qquad (15.2)$$

15.5 Use of average breeders

For species spawning only once and for which there are no reliable methods for cryopreservation or the costs connected to cryopreservation prohibit application of repeat matings as described above, a control involving use of average breeders may be considered (Gjedrem et al., 1997). A number of average breeders (i.e. breeding candidates with average estimated breeding values) are used to produce a pooled group of offspring, which serve as control in the following generation. The rationale behind this approach is that the offspring produced from the average breeders are assumed to represent the average genetic level in the parent generation. The selection response can be estimated based on the difference between the average performance of offspring from the selected and the average breeders.

Figure 15.1 illustrates how this method can be applied for estimation of genetic gains for growth rate based on individual selection. At the same time as the largest fish are selected for production of improved offspring, a number of average sized breeders is used to produce the animals serving as controls. Pooled groups of offspring from selected and average breeders are reared separately until a random sample from each group can be tagged and subsequently stocked and reared in the same production unit. When the fish reach the target size, the genetic gain for growth can be estimated as the difference in average weight between the two groups.

In many practical situations, use of a control group using average breeders is appealing. This method can easily be facilitated for most species at reasonable cost, and the animals belonging to the control group are expected to perform better than control animals produced by methods involving control lines, in particular when selection has been conducted for several generations. On the other hand, the usefulness of this approach may be limited when selection is simultaneously done for many traits, and in particular if the breeding goal includes traits showing antagonistic genetic relationships. In that case, selection of average breeders based on the aggregate breeding value (i.e. net merit) may tend to pick breeders with breeding values that deviate strongly from the population mean for each of the unfavorably genetically correlated traits and thus disqualify them as representative average breeders for the individual traits under selection. Under such circumstances it might be considered more useful to produce a separate control group for each individual trait in the breeding goal, or for those of main

interest if the number of traits under selection is high. Production and maintenance of several average breeder control groups may, however, considerably increase the costs involved. Furthermore, the number of pairs of average breeders used for production of each control group must be sufficiently high and their relative contribution of animals in the mixed control group need to be balanced in order to avoid sampling variance that may bias the selection response estimate obtained by this method.

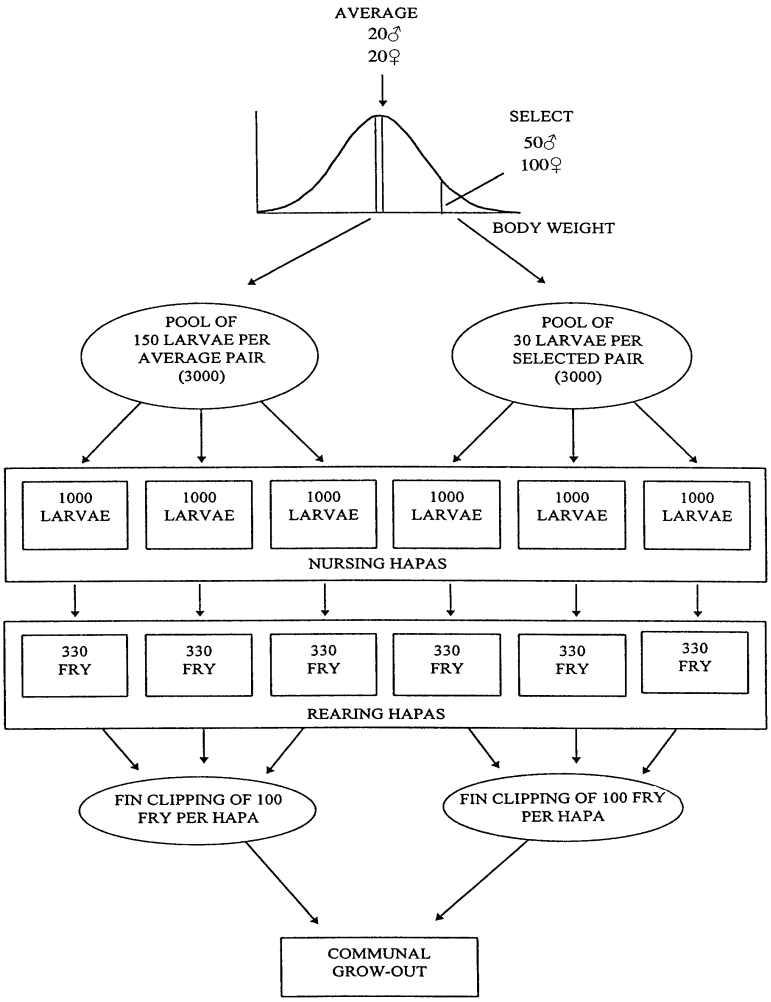


Figure 15.1 Control of genetic gain. Comparison of progeny of selected breeders with progeny of average breeders. Reproduced from Gjedrem et al. (1997) by permission of WorldFish Center.

Examples on the use of average breeders for estimation of genetic gain for growth rate in carp and tilapia are given by Mahapatra et al. (2004) and Bentsen et al. (2004), respectively.

15.6 Genetic trend analyses

Estimates of selection response can also be obtained by a statistical approach. The estimation of genetic parameters (Chapter 9) and prediction of breeding values (Chapter 13) in improvement schemes for livestock worldwide is now largely dominated by mixed model methodology introduced by Henderson (1975). These methods are also rapidly gaining grounds in experimental and applied improvement programs for aquaculture species. Assuming that phenotypic and genetic variances are known, BLUP (best linear unbiased prediction) procedures allow simultaneous estimation of fixed effects and breeding values. Although the assumption of known variances seldom holds, it is demonstrated that good approximations of the genetic and phenotypic variances can be provided by restricted maximum likelihood (REML).

Mixed model methodology also provides a valuable tool for accurate estimation of selection responses through genetic trend analyses. Of particular interest in this regard is the animal model, representing the most general and yet advanced BLUP model. The animal model provides estimated breeding values for all animals included in the analysis, and uses all the relationships between these animals. If data from more than one generation is analyzed, the genetic trend reflecting selection response adjusted for environmental trends in the data can be estimated by regressing the average breeding values within generation on generation number. Unbiased estimation of selection responses by the animal model, however, requires that BLUP accounts for all performance and pedigree information traced back to the unselected and non-inbred base population. Similarly, accurate separation of genetic and environmental effects requires sufficient genetic ties between fixed effect classes within year class as well as between generations (Sorensen and Kennedy, 1984a).

It should be underlined that since a genetic trend analysis is of pure statistical nature and does not involve any external reference point; the resulting estimates of response are largely determined by the genetic parameters used in the analysis. Unbiased genetic parameter estimates should therefore be used with this approach.

For a detailed account on application of mixed model technology for estimation of genetic gain in fish, see Gall and Baker (2002).

15.7 Prediction of response based on selection paths

In many situations (e.g. for optimization of selection schemes) it may be useful to quantify the relative contributions to the overall selection response originating from the male and female side, respectively. Annual genetic gains can be predicted as the genetic superiority of sires and dams divided by the generation interval, as discussed by

VanVleck (1987) for individual selection. Let the average breeding values for sires be S and for dams D. Under individual selection, the genetic superiority of sires will be:

$$\Delta S = i_S \, h^2 \, \sigma_P \tag{15.3}$$

where i_S is average selection intensity for sires, and h^2 and σ_P is the heritability and the phenotypic standard deviation, respectively, for the trait in question. Similarly the genetic superiority of dams is predicted by:

$$\Delta D = i_D \, h^2 \, \sigma_P \tag{15.4}$$

where i_D is the average selection intensity for dams. The expected genetic gain per year in the progenies will be:

$$\Delta G = (\Delta S + \Delta D)/(L_S + L_D) \tag{15.5}$$

where L_S and L_D are generation interval in years for sires and dams, respectively. When $L_S = L_D = L$ the estimate of genetic gain can be expressed as:

$$\Delta G = (\Delta S + \Delta D)/2L \tag{15.6}$$

The values of ΔS and ΔD tell how much each of these two paths contributes to the total genetic gain in the population.

Similarly, the selection paths between parents and progeny can be extended to grand parents and consider sire of sire (SS), dam of sire (DS), sire of dam (SD) and dam of dam (DD). Each of these four paths may have different generation interval (L_{SS}, L_{DS}, L_{SD} and L_{DD}).

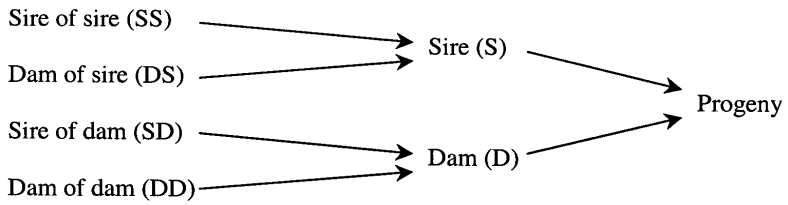


Figure 15.2. There are four paths from grand parents through parents to progeny.

The genetic superiority of sires to sires will be:

$$\Delta SS = i_{SS} \, h^2 \, \sigma_P \tag{15.7}$$

where i_{SS} is the selection intensity of sires to sires, and similar for the other paths. The expected genetic gain can be written as follows:

$$\Delta G = (\Delta SS + \Delta DS + \Delta SD + \Delta DD)/(L_{SS} + L_{SD} + L_{DS} + L_{DD}) \quad (15.8)$$

The values of ΔSS, ΔDS, ΔSD and ΔDD tell how much each of these paths contributes to the total genetic gain in the population.

This method can also be extended to cover more general breeding schemes involving family or combined family and within- family selection.

16. BREEDING PLANS

TRYGVE GJEDREM

16.1 Introduction

There is much to be gained in increased productivity from the domestication and development of genetically improved stocks in aquaculture. But even the simplest breeding program needs investment and requires input of labour and personal well educated in quantitative genetics, in order to be successful. It is therefore difficult for small farms to carry through breeding programs in a profitable way in isolation. In farm animals, it has been common practise that farmers form national cooperatives to run breeding programs, in order to serve their members with improved stock. Another approach is that private companies run breeding programs to service international markets.

An essential part of the breeding program is development of strategies for testing and programs for estimating breeding values. In pigs and dairy cattle, these strategies are well advanced. In order to pay off, a breeding program covering the economic important traits should serve a relatively large population of animals.

Before the start of a breeding program a detailed plan must be made taking into account the reproductive biology as well as the fecundity of the species in question. Further it is a necessity that phenotypic and genetic parameters, estimated with relatively high accuracy are available for the traits and species in question. In fact, as soon as aquaculture production focuses on new species, it is important to carry through breeding research projects in order to estimate phenotypic and genetic parameters and thus lay the theoretical basis for breeding plans.

Based on discussions in previous chapters, the different elements in a breeding program will now be summarised and discussed.

16.2 Fish and shellfish as targets for breeding programs as compared with farm animals

Farmed fish and shellfish species have not gone through as many generations as the farmed species of mammals and birds have. The majority of the genetic material used in fish and shellfish farming has not been entirely separated from the wild populations to form specialised domesticated breeds. Wild broodstock have often been used as a regular part of fish farming systems, either continuously or to "refresh" the cultivated stocks. Furthermore, selection has rarely been successfully practised in fish and shellfish probably because of lack of controlled mating and reproduction and lack of appropriate and cheap tags. The possibility of keeping pedigree records has therefore been limited.

T. Gjedrem (ed.), Selection and Breeding Programs in Aquaculture, 251–277.
© 2005 *Springer. Printed in the Netherlands.*

Aquatic species have several advantages as the target of breeding programs compared to farm animals, notably:

- They have very high fertility compared with farm animals, where only one or a few eggs per female are common
- Most of the species have external fertilisation, which usually can be controlled artificially. This allows great flexibility in mating schemes. The semen or eggs collected at each spawning can be divided in many groups and mated to produce full and half-sib groups as well as different crosses
- Compared with most farm animals, fish and shellfish show a remarkable potential for inter-specific hybridisation
- In connection with fertilisation it is relatively easy to manipulate chromosome number to produce haploids, triploids and tetraploids in fish and shellfish
- Production and rearing of breeding animals is very cheap compared with farm animals

But aquatic species have also some disadvantages as the targets of breeding programs compared with farm animals:

- At hatching alevins or larvae are tiny creatures, and are impossible to tag or mark. Therefore in order to keep pedigree records each full-sib family must be reared separately until they reach a size big enough for tagging. This size may be several grams. This is a problem and will usually make breeding experiments as well as breeding programs rather expensive and will introduce environmental variation between families. Modern technology by applying microsatellites for parental identification may reduce this problem
- Most of the farmed fish and shellfish are small animals with low individual economic value. Many farmers and researchers therefore believe that it is too costly to develop sophisticated and efficient breeding programs for these animals
- In a closed population, high fecundity can easily lead to a rapid build up of inbreeding if special precautions are not taken

16.3 Consequences of level of fecundity

Since the theoretical basis for quantitative genetics was developed in the middle of 1930's, modern breeding plans have been created for farm animals and plants. In plants, non-additive genetic variance is relatively large for important traits and therefore crossbreeding to produce hybrid vigour has been a central element of the breeding plans. In cattle, pigs, sheep and poultry, the economic traits show primarily additive genetic variance, and therefore use of pure-breeding with selection has played a central role in

breeding programs. In aquatic animals the non-additive genetic variance has been shown to be relatively low compared to the additive component (Chapter 5), and therefore a breeding program for fish and shellfish will be more similar to one for farm animals than for plants.

In fish and shellfish the fecundity is very high in all farmed species. Tilapia is one of the species with lowest fecundity giving 300 to 1500 eggs per spawn. However, under optimal rearing conditions they will spawn regularly and thus produce many eggs per year. The consequence of the high fecundity in fish and shellfish is that the breeding program can be centralised which will reduce the running cost compared with field tests. High fecundity makes very intensive selection possible.

Several authors have discussed the processes of developing breeding programs for aquaculture species, in light of their high fecundity (Gjedrem, 1983a, 1985, 1992; Shultz, 1986; Gall, 1990; Bentsen, 1990; Refstie, 1990). The main elements in breeding programs in farm animals are recommended, starting with the formulation of breeding goals.

16.4 Breeding goal

Most animal breeding programs have so far focused on cumulative short-term genetic change of production traits, because the optimisation of the breeding program has to a very large extent been based on market economic values. Sustainable genetic improvement by animal breeding on the other hand is a long-term and complex process. Breeding programs to achieve this goal need to focus on long-term biological, ecological, and sociological solutions. Precautions should be taken to avoid unwanted side effects as are often seen in breeding programs for agricultural livestock, examples are leg weakness in pigs and poultry, and reproductive traits in cattle. Careful monitoring of possible correlated responses is needed. More basic knowledge about welfare and behavioural needs of fish may also be required.

In order to achieve a holistic approach for animal breeding, Olesen et al. (2000) suggested a procedure where:

1. The ethical aspects and priorities should be made clear
2. The system should be defined with respect to limits and structure, resource efficiency, environment, economics and social effects
3. Indicators should be defined that measure or characterize the above ethical priorities and critical effects of the production system
4. Performance traits and characters that are important or critical to meet these criteria or objectives should be identified and weighted

The system must be optimized according to ethical priorities, and the objective of animal genetic improvement is to fit animals' traits according to this. So far, weighing of traits has mainly been dependent on economic values and frequency of expression or

mean of the trait (e.g. mean growth rate and survival rate). The methodology to weigh the traits with respect to resource efficiency and economy is well developed.

Olesen et al. (2000) suggested to split the trait values in the aggregate genotype into non-market values (NV) and pure market economic values (MeV). This gives the following breeding goal or aggregate genotype (H, considering two traits Y_1 and Y_2):

$$H = [NV_1 \times Y_1 + MeV_1 \times Y_1] + [NV_2 \times Y_2 + MeV_2 \times Y_2] \quad (16.1)$$

Correspondingly, a genetic gain of non-market value will be obtained in addition to a gain in the economic traits. The value of the non-market gain is $NV_1 \times \Delta G_1 + NV_2 \times \Delta G_2$ and likewise for the market economic gain ($MeV_1 \times \Delta G_1 + MeV_2 \times \Delta G_2$). The total genetic gain is a sum of the non-market genetic gain and market economic genetic gain.

As shown, a trait may have both non-market and market value. Reduced disease frequency increases value through improved animal welfare and increases market economic value through reduced economic costs of treatments, reduced yield and increased market acceptance of the product. When fish production relies on long-distance energy inefficient transport of feeds, feed efficiency should be emphasized through both market and non-market values. However, a trait may have only non-market value or only market economic value. The resulting non-market and market genetic gains provide an opportunity to evaluate the breeding programs in a holistic perspective, where both social, cultural (including subjective values), ecological, and economic objectives and effects can be taken into account. Also, there are already developed methodologies for quantifying values of for example environmental services and animal welfare, where subjective opinions may be strong and vary considerably among individuals and cultures (Braden and Kolstad, 1991; Freeman, 1993; Smith, 1993; Bennett, 1996). Olesen et al. (1999) considered these methods, and concluded that they can be applied for estimating non-market values of traits included in animal breeding goals. However, further development and adjustment of these methods to animal breeding will be needed, as there are few examples of implementations of this approach to date.

The definition of a breeding goal (breeding objective) is the important part of a breeding program. It is crucial that the breeding companies and farmers agree on the general goal, so that everybody works together in the same direction within one breeding program. The general aim is to develop animals with high productivity and quality products. Within this general breeding goal it is possible for the breeding unit to offer breeders selected animals for certain traits or specialized lines.

Industry and the consumers should establish the breeding goal after consulting with processors and exporters. Since the request for product quality may vary from one market to another, differentiation in quality and sizes for different markets may be needed, which could be an argument for several specialized breeding goals. Dividing a population into several sub populations with different breeding goals, will increase the cost and usually result in higher cost of the breeding activity per animal. An alternative

is to produce offspring specialised in certain traits, such as increased resistance to certain diseases. This could be obtained by intensively selecting broodstock with high breeding value for the specific trait in question.

It is preferable that traits in the breeding goal will be of importance for the foreseeable future. Some traits like growth rate, feed conversion efficiency and disease resistance, are quite stable and the breeding goal for these traits is not likely to change in the near future, which makes them easier to work with. This must be taken into account when breeding goals are discussed, since response to breeding and selection are obtained in future generations, when the first and later generations of progenies are marketed, not when the breeders are selected.

For a trait to be included in the breeding goal, the following prerequisites must hold:

- The trait must be of economic and ethical importance
- The trait must show genetic variance
- It must be possible to measure the trait at a reasonable cost

Definition of the breeding goal becomes more and more important for an effective multitrait selection program, especially when traits in the breeding goal are highly negatively correlated and have different economic values. Biological knowledge of the genetic background of the gross composite breeding goal traits and the mutual relations may lead to better definition of breeding goals and selection criteria.

16.4.1 GROWTH RATE

For meat and flesh producing animals growth rate is the most important economic trait. High growth rate increases production turnover and a fast growing fish will reach a higher body weight before the onset of sexual maturation. A fast growing animal will need less feed for it's maintenance requirement simply because it reaches market size faster than the slower growing animal. Growth rate is easy and cheap to record and can be measured exactly either as body weight or as body length. Body weight has been shown to be a heritable trait and highly variable (Table 5.1 and 5.7 and Figure 10.1). Body weight of fish usually has a coefficient of variation of 20-30 % while shrimp and oysters seem to have a lower coefficient of variation. In Figure 16.1 variation among shrimp families *(P. vannamei)* in harvest weight reared in two environments are shown. The correlation between family averages in these two environments was 0.87. Genetic gain in growth rate can be harvested by producing a larger animal or by producing animals at the same size as before but in a shorter time.

16.4.2 FEED CONVERSION EFFICIENCY

In fish and shellfish species the feeds represent a large portion of total production cost. Feed conversion efficiency is therefore of great economic importance. However, it is difficult and costly to record the amount of feed consumed by individuals, or even for families of aquatic organisms. It is particularly difficult to record food not eaten after the animal is fed as some of the nutrients will dissolve in the water. One possible alternative

is to measure parts of the feed consumption, or try to find indirect measures which are correlated to feed conversion efficiency.

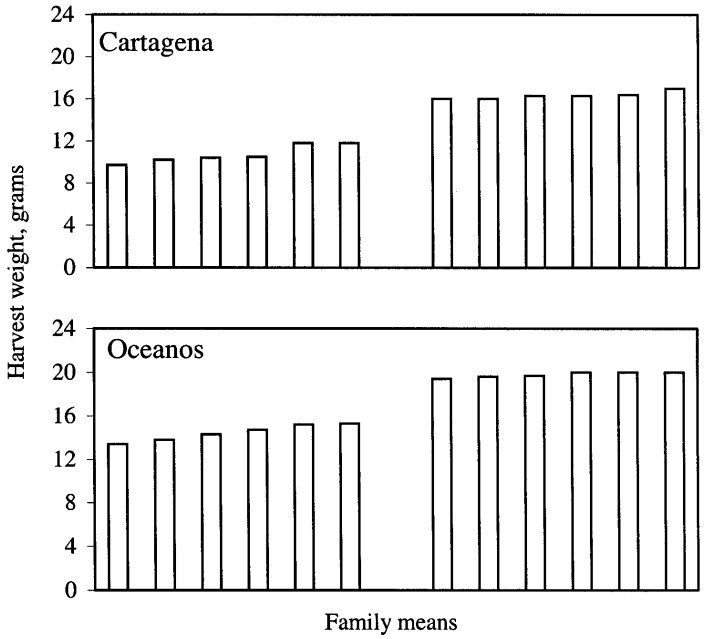


Figure 16.1. Ranking of shrimp full-sib families for harvest weight reared in two different environments, Cartagena and Oceanos. Reproduced from Rye, et al. (1999) by permission of AKVAFORSK and CENIACUA.

In farm animals relatively high genetic correlations are found between growth rate and feed conversion efficiency (Andersen, 1977, r_G = -0.95 in cattle; Vangen, 1984, r_G = -0.93 in pigs; Crawford, 1990, r_G = -0.20 to -0.80 in broiler). Kinghorn (1981) found a high positive genetic correlation in young rainbow trout between growth rate and gross feed conversion (Growth/feed consumed) and a negative genetic correlation between growth rate and net feed conversion (Growth/feed consumed less the maintenance ration). In rainbow trout Gjøen et al. (1993) estimated the genetic correlation between feed consumption and specific growth rate to be r_G = -0.78. If these high genetic correlations between growth rate and feed conversion efficiency reported above are true for fish and shellfish, selection for increased growth rate will result in a considerable correlated response in feed conversion efficiency.

Results presented by Thodesen (1999a) indicate a genetic variation in feed efficiency ratio (FER) in Atlantic salmon, opening up the possibility to improve feed efficiency through direct selection, which is further discussed in Chapter 7.

16.4.3 DISEASE RESISTANCE

Diseases represent one of the major problems in fish and in shellfish. In shrimp various virus diseases have lately resulted in the loss of most or all of the farmed populations in several countries. The possibility of the spread of water borne infectious diseases significantly increases the risk of farming fish and shellfish. Disease is observed as change in behaviour, reduced appetite, lesions and finally dead animals. In practical farming it is not possible to identify and record infected survivors with reduced performance but only dead animals. Survival is a complex trait and the causes of death may differ from one farm to another, and from one year to another. This makes survival difficult to use as breeding goal. If necessary survival rate can be recorded during the period when the families are reared separately. In later stages of production it is more difficult but possible if animals are tagged with electronic tags. Records of survival show low heritability in general (Table 5.7).

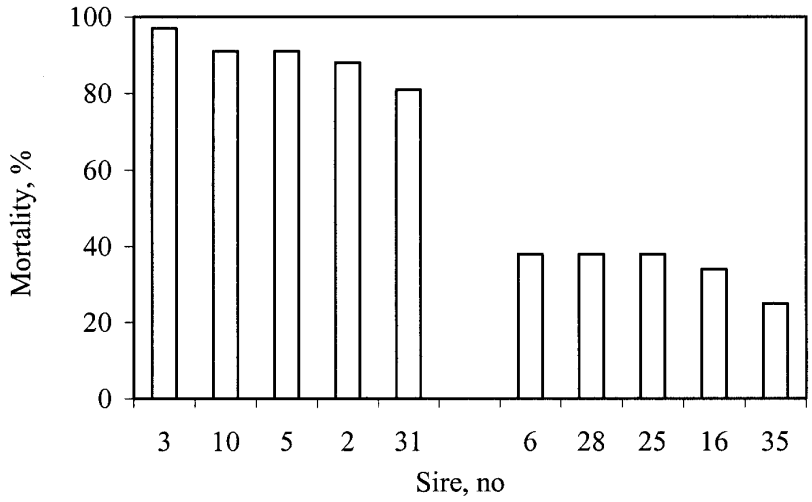


Figure 16.2 Variation in mortality between full-sib families of Artlantic salmon challenged by furunculosis bacteria. Reproduced from Gjedrem et al. (1991) by permission of Elsevier.

Measuring survival to specific diseases may give more scope for improvement. Recent research has shown relatively large genetic variation in disease resistance by applying challenge tests for specific diseases. In Figure 16.2 variation in frequency of mortality between full-sib groups due to a challenge test are shown for Atlantic salmon. Atlantic salmon were tested for resistance against furunculosis bacteria infection.

Since direct measurement of disease resistance is rather difficult and since challenge test is rather expensive, one possibility is to try and find markers which are highly correlated to disease resistance. Immunological parameters have been proposed as candidates for correlated traits in order to obtain correlated response in disease resistance, but as

reported in chapter 7 none of the genetic correlations between immunological traits and disease resistance exceeded ± 0.37. For indirect selection to be efficient the genetic correlation between the genetic marker and disease resistance should be >± 0.5.

16.4.4 AGE AT SEXUAL MATURATION

Early sexual maturation is a disadvantage in some species because feed is converted into gonads instead of into meat. Sexual maturity will generally reduce growth rate, reduce meat quality and increase mortality. The breeding goal should therefore be that animals should reach market size before sexual maturity.

In some species it is desirable to select against early sexual maturation. Salmon and tilapia are examples. Most of the production of rainbow trout is for portion size fish, so there is no problem with early maturation. In Atlantic halibut males approach sexual maturation very early (at about 2 kg) and the growth rate is markedly reduced. On the contrary some species reach sexual maturity much later than they reach market size of which grass carp, silver carp, Bighead carp and milkfish are examples. The breeding goal and the importance of this trait may change over time in populations where growth rate is improved by selective breeding.

Considerable genetic variation has been demonstrated for age at early maturation (Longalong et al., 1999) and some estimates of heritability are given in Table 5.7. However, the possibility for genetic improvement of this trait is to a large extent dependent on the frequency of maturing fish at a given age. Age of sexual maturation is recorded as mature or not mature at a defined age and is called an all or none trait. Family selection is recommended for such traits particularly when the frequency of the trait is low, 10-20 %.

16.4.5 QUALITY TRAITS

Demand for high quality is particularly important for food products and is therefore one of the central goals in fish and shellfish production. It is expected that quality traits will become more and more important as the industries develop. However, the demand for quality will vary considerably from one species to another, from one market to another, and will also vary over time. The main problem with quality traits from a breeding point of view is that most of them can not be measured on live breeding candidates, but only on dead relatives. Family selection must therefore be applied since individual selection is inefficient (Gjedrem, 1997a). Estimates of genetic parameters for quality traits are given in Table 5.7.

Fish quality is usually a question of size, meatiness, fat percentage, fat distribution, flesh colour, fillet yield, structure, firmness, taste, shape of body, dressing percentage and so on. For shellfish more emphasis is put on size of mussel, tail percentage, taste and colour. Furthermore some of the traits are rather difficult to measure or judge, such as meatiness, taste, structure and firmness.

Some fish species like carp, silverbarb and sea bream have intramuscular bones. These small bones are a problem because they are difficult to remove and the consumers do not want them. There have only been a few studies concerning genetic variation in intramuscular bones in fish. In one population of common carp large phenotypic variation was found (Sengbusch, 1963, 1967; Sengbusch and Meske, 1967; Meske, 1968). Kossman (1972) found genetic variation in intramuscular bones in common carp, while Moav and Finkel (1975) did not find genetic variation for this trait. Thus the magnitude of genetic variation in this important trait is unclear.

Condition factor is commonly used as a measure of form of fish and is a ratio of body weight (g) and body length cubed (cm^3). Dressing % is a measure of how much of the body may be considered as waste. This includes blood, intestine, fat in and around the intestine and head. In general a high dressing % means that a large part of the carcass is edible. However, one should not try to reduce the size of the intestine, because these organs are of vital importance in metabolism and feed digestion (Gjerde and Schaeffer, 1989).

16.4.6 FECUNDITY

Fish and shellfish have very high fecundity (see Table 8.1) and as body weight of broodstock will increase by selection number of eggs and amount of milt will increase since gonad size increases with body size. It should therefore not be necessary to select for increased number of eggs in a breeding programme.

It is, however, important to record fecundity traits, and particularly rate of fertilization, mortality of eggs and alevin, deformity of alevin and other egg and sperm quality traits in order to study whether or not these important traits are altered over time. Unfavourable changes in such fitness traits should be a warning that the directional selection is having adverse affects.

16.4.7 RECAPTURE FREQUENCY

Sea ranching is in particular practised for anadromeous species and recapture frequency is the most important economic trait (Isaksson, 1988). Recapture frequency is a complex trait, of which survival is an important component. In Atlantic salmon recapture frequency has been shown to be heritable (Carlin, 1969; Jonasson et al., 1997). Consequently it should be possible to increase recapture frequency by applying family selection.

16.4.8 BEHAVIOUR

Behaviour of the animals in aquaculture farms will influence the production. Wild, undomesticated animals tend to be restless and nervous. Feed will easily be wasted and the feed conversion rate is likely to be low. Non-domesticated animals will experience a high level of stress and will not thrive. The high level of stress will reduce their resistance against diseases. In order to reduce these problems selection is a possibility.

Tanck et al. (2000) estimated a heritability of 0.60 for stress-related cortisol increase in five months old carp. The stress was obtained by a cold shock to 9^0C. The problem is, however, that behaviour is difficult and laborious to record both on an individual and family basis. However, experience shows that selecting fish for high growth rate and other traits in several generations, will favour animals, which are better adapted to farm environment, calmer, less afraid of people and less aggressive. As long as we do not have appropriate and inexpensive methods to measure behavioural traits accurately, it is difficult to include them in the breeding goal.

16.4.9 SUMMARY

It is not possible to define a general breeding goal since it will vary from species to species and among production systems, but the following general ranking of traits according to importance is often the case:

Aquaculture farming:
1. Growth rate
2. Disease resistance
3. Quality traits
4. Age at sexual maturation (for some species)

Sea ranching:
1.Return frequency
2. Growth rate

16.5 Forming a base population

When developing a breeding program for aquaculture species it is important to start with a broad genetic base. There are examples indicating that mass selection for improved growth rate in experimental populations of carp and tilapia, may have failed because of narrow genetic variation in the base population (Moav and Wohlfarth, 1973, 1976; Hulata et al., 1986; Teichert-Coddington and Smitterman, 1988; Huang and Liao, 1990). Even if some response was observed for downward selection, Moav and Wolfarth (1976) suggested that genetic bottlenecks and high levels of inbreeding in such closed, experimental populations might have reduced the genetic variation significantly.

The genetic variability in the base population may be secured by forming a synthetic population (Skjervold, 1982). The first step in forming a synthetic base population should be to compare productivity in available populations. A base population should combine characteristics of the subpopulations. Bondari (1983) derived a genetic base for channel catfish by crossing six different cultured stocks and obtained significant response to selection for growth rate. The Norwegian breeding program for Atlantic salmon was started by collecting and testing breeding candidates from 41 wild river strains (Gunnes and Gjedrem, 1978). The base for the synthetic tilapia population used in the selection program, GIFT (Genetic Improved Farm Tilapia), was an 8x8 diallel cross between eight tilapia strains (Eknath et al., 1993). In India six stocks were crossed to form a base population of rohu (Reddy et al., 2000).

Even if cultured strains are available, wild populations may still be possible contributors to a synthetic population because the genetic differences between wild and cultured stocks may be moderate (Bentsen and Gjerde, 1994). Selecting the best individuals across populations should form the founder stock of a synthetic population. Some minimum level of representation from the tested populations may be demanded, to ensure the genetic variability of the synthetic population (Bentsen, 1990).

The choice of parent strains in a hybrid program is determined by the non-additive genetic interaction in the crossbred offspring. Strains, which combine best in crossbreeding, can only be found by crossing available strains. The hybrid performance should not only exceed the best parent strain, but also all other available purebred stocks. In general, crossbred offspring of highly inbred strains may express considerable heterosis, while heterogeneous parent strains normally produce less hetaerists. This picture seems to be valid also for fish crossbreeding (Gjedrem, 1985; Dunham, 1986; Bakos, 1987). A hybrid program without selection for improved additive performance will be uncompetitive with a long-term selective breeding program.

16.6 Breeding strategies

Since inbreeding in general is harmful, it should be avoided in any breeding program except when line crossing is applied. Two breeding methods have practical use: crossbreeding and purebreeding. The choice of which of these breeding methods, or which combination of the two, depends on what kind of genetic variation is present in the traits in the breeding objective.

Crossbreeding should be used if non-additive genetic variation is large, while purebreeding combined with selection should be used to exploit additive genetic variation. It is not always easy to decide which breeding method should be chosen. Gjedrem (1985) concluded that if there is additive genetic variation, one should always use purebreeding and apply selection. If the component of non-additive genetic variance is considerable, selection should be combined with crossbreeding. According to Gjedrem and Fimland (1995) crossbreeding should as a rule be used in a breeding program when the magnitude of heterosis is larger than the expected genetic gain per generation in a purebreeding program with selection. It must be kept in mind that crossbreeding will increase the cost of a breeding program considerably, because at least two lines or populations must be used and tested. Therefore a combined pure and crossbreeding program requires testing facilities which are about twice as large as a purebreeding program.

The results from crossbreeding experiments vary considerably, Chapter 10, and only a few show high value of heterosis. It is therefore appropriate to conclude that purebreeding with selection should be used in most breeding programs for fish and shellfish. However, new results may modify this conclusion and justify more crossbreeding programs.

16.7 Selection methods

So far additive genetic variation has been found for all traits of economic importance in all species studied, Table 5.7. Therefore selection will be the main and central method to be used in breeding programs for fish and shellfish. As concluded in Chapter 11 individual and family selection is of particularly interest to be used in breeding programs for fish and shellfish while progeny testing is not so valuable since the generation interval will be increased and usually doubled. However, progeny testing may be of considerable interest for species spawning over several years.

Individual selection is the most widely used method of selection. It is easy to practise, the cost of testing and recording to estimate breeding values is low compared to other methods of selection and it does not extend the interval between generations. However, individual selection has limitations. The efficiency is low for traits with low heritability; it can hardly be used for all- or- none traits such as mortality and age at sexual maturation, and it can not be used for flesh quality traits. In breeding programs for aquatic animals, individual selection is usually limited to growth rate.

For other traits family selection should be used since within-family selection has shown low efficiency (Gall and Huang, 1988a; Gjedrem, 1985). Family selection is relatively efficient compared to individual selection for traits with low heritability. It should be applied for all-or-none traits and it must be used for meat quality traits. Family selection does not lengthen the generation interval as progeny testing does. Since it is impossible to tag newly hatched larvae or fry, each family must be reared in separate tanks for the first period of their lives, before they can be tagged. This type of rearing, however, introduces some environmental effects common to families, which are called tank effect. Refstie and Steine (1978) estimated the tank effect on body weight for Atlantic salmon smolts to 5 % of total phenotypic variance while Aulstad et al. (1972) estimated the tank effect in 150 and 280 days weight of rainbow trout to 6.2 % and 4.3 %, respectively. Tagging of full-sib families makes it possible to keep pedigree records over generations and avoid mating of relatives.

In general, it can be concluded that individual selection should be used for growth rate in all species and family selection for the other traits included in the breeding goal. When family selection is applied a combined family and within family selection should be used for growth rate.

16.8 Genotype - environment interaction.

When planning a breeding program, it is important to know if genotype - environment interaction is present. With no interaction, the breeding plan can concentrate on the best strain or combine the best strains into a synthetic population. On the other hand, if interaction is significant and account for a relatively large part of the total variance, the response to selection will be reduced and, consequently, it may be desirable to develop strains for different environments.

So far no genotype - environment interaction of importance has been shown in shellfish. In rainbow trout, Atlantic salmon, tilapia and selected lines of channel catfish, negligible genotype - environmental interaction has been demonstrated (Refstie and Steine, 1978; Gunnes and Gjedrem, 1981; Dunham, 1987; Eknath et al., 1993). In common carp, crossbred channel catfish and rainbow trout reared in varying salinity and prairie lakes, considerable genotype - environmental interaction has been shown (Moav, 1976; Dunham, 1987; Sylven et al., 1991; Ayles and Baker, 1983).

Thus it appears that one can not draw a general conclusion about how many strains or breeds should be used in a breeding program. That question must be investigated for each species and each location. It is, however, likely that genotype - environment interaction will increase with the variability of environmental conditions. Genotype–environment interaction is further discussed in Chapter 14.

16.9 Identification of animals

There is a great need for cheap methods of tagging small and young fish and shellfish in order to identify the animals and to keep pedigree records. Refstie and Aulstad (1975) list the following requirements for satisfying marking methods for fish:

- The method should be applicable for small animals
- The tag should not influence the growth rate of the animal
- The method should be inexpensive
- The method of tagging should require little labour
- The mark should be readable past the time for recording

These requirements are also valid for shellfish.

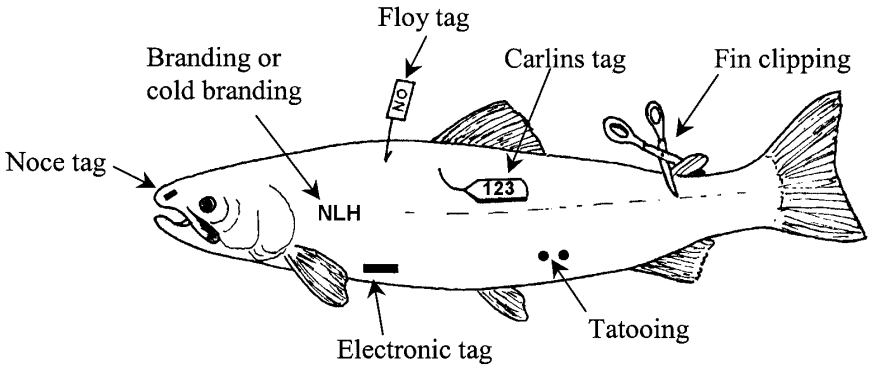


Figure 16.3 Different types of tags for identification of fish.

A great number of different marking methods are described in the literature. Many types of physical metal or plastic tags with numbers and addresses on are fastened to fins, jaw,

tail and gill cover by wire or string which may be elastic, Figure 16.3. Some metal tags are shot into the nose (nose tag) or inserted into the body cavity (pit-tag). Cold-branding, branding and tattooing together with fin clipping are commonly practised. There are really no completely satisfactory tags available for fish and shellfish. The most promising tag is electronic tags (pit tag), which may be inserted into the body cavity. The electronic tag may have large numbers and can be red by a special device at any time whether the fish is alive or dead. However, the pit-tags on the market to day are expensive. For shrimp an internal colour tag can be used on juveniles as small as 1 g (Godin et al., 1996) and plastic tags can be attached to the eyestalk of larger shrimp.

Refstie and Aulstad (1975) describe a system for tagging using a combination of cold branding and fin clipping. This method of tagging was used in the Norwegian breeding programme for Atlantic salmon and rainbow trout from 1973 until 2000 but this method has strong limitations as only a limited number of families can be tagged, the tag can not be read on dead fish and some fish can not be identified at harvest.

The development of microsatellite DNA profiling techniques for family identification purposes represents a milestone in selective breeding of aquatic species. These techniques allows different family groups to be kept in a common tank from fertilization onwards, substantially reducing the costs involved and eliminating problems related to common environmental effects. It also enables larger numbers of families to be tested and thus facilitates the use of higher selection intensities without rapid accumulation of inbreeding. However, the readings of DNA sequences are very expensive.

A genetic identification system will reduce the need for physical tagging, since only surviving individuals need to be identified. This will particularly be valuable in species where culture is undertaken only for part of the life cycle and the harvested fish constitute a rather small proportion of the total number released. The wire tags presently used in sea ranching programs (Jonasson, 1994) can not be read on live individuals and imposes severe constraint on the use and mating of the best individuals.

16.10 Closed life cycle and controlled mating

In an efficient breeding program it is a prerequisite that the life cycle of the species involved is closed. This requires that it is possible to reproduce the animals and produce broodstock artificially, which mean that wild animals are not used as broodstock. As soon as the life cycle is closed the process of domestication will start. It is very important that the animals become adapted to the practical farming system which is used. Wild animals will never thrive in captivity and they will live under continuously stress while domesticated animals will become more and more adapted to captivity.

Quite a number of freshwater fish can be artificially reproduced like carps, salmonids, tilapia, catfishes, silverbarb, pike, sturgeons and so on, while the reproduction cycle is only known for few number of marine fish species. Marine species that can be artificially reproduced are flounder, turbot, halibut, cod, sea bass, sea bream, yellowtail,

mahimahi, mullet and milkfish. For most of the farmed species of shellfish the reproduction cycle is at least partly known.

For carp species it was a break through when induced spawning with gonadotropin from pituitary extracts of physiologically remote fish were discovered. Carp will only spawn naturally in running waters. A Brazilian biologist was the first to use induced spawning in 1934 and the method was used in China from 1954 and from then on it become widely applied. Induced spawning is also successfully used in other species like seabass, sea bream and catfish.

In a breeding program it is important to be able to control the mating. This is a necessary to avoid mating of close relatives and thus control the build up of inbreeding in the population.

16.11 Mating design

Most fish and shellfish species have external fertilisation, which in combination with their high fecundity opens large flexibilities and possibilities in selection experiments and breeding programs. It is therefore possible to practise a variety of mating design such as: factorial, nested or hierarchical, and combination of them. The advantage and limitations of these mating designs has been discussed in detail in Chapter 12 together with possibilities for estimation of phenotypic and genetic parameters and breeding values.

16.12 Testing strategies

The purpose of testing animals is to obtain information about their genotype or breeding value. It is therefore important to keep the animals under uniform conditions in order to reduce the environmental component of variation as much as possible. The examples of moderate success by mass selection both historically and in recent years, may have been partly caused by large uncontrolled systematic as well as random environmental variation and by uncontrolled inbreeding. Social competition may be another systematic source of variation in growth rate. Even if competing ability is assumed to be partially genetic, it is not expected that the correlated response in competing ability will improve growth rate (Kinghorn, 1983a; Doyle and Talbot, 1986a). Social competition may also magnify age dependent variation.

Several strategies have been suggested to avoid these problems, such as age standardisation (Tave and Smitherman, 1980), size standardisation (Doyle and Talbot, 1986b) and within family selection procedures (Uraiwan and Doyle, 1986). However, the standard strategy in other farm animals has been to test breeding candidates in a common environment without severe restrictions on resources affecting performance, and to maintain individual life history records (Bentsen, 1990). The testing environment should be similar to a representative, commercial farming environment. This strategy

has successfully applied to fish (Gjedrem, 1979b; Bondari, 1983; Hershberger et al., 1990) and shrimp (Fjalestad et al., 1997; Hetzel et al., 2000). Standardisation of environmental conditions and facilities during the test period for the genetic groups under comparison should be given high priority, like water temperature and quality, light intensity and photoperiod, amount, content, and quality of feed, animal density and so on (Bentsen, 1990).

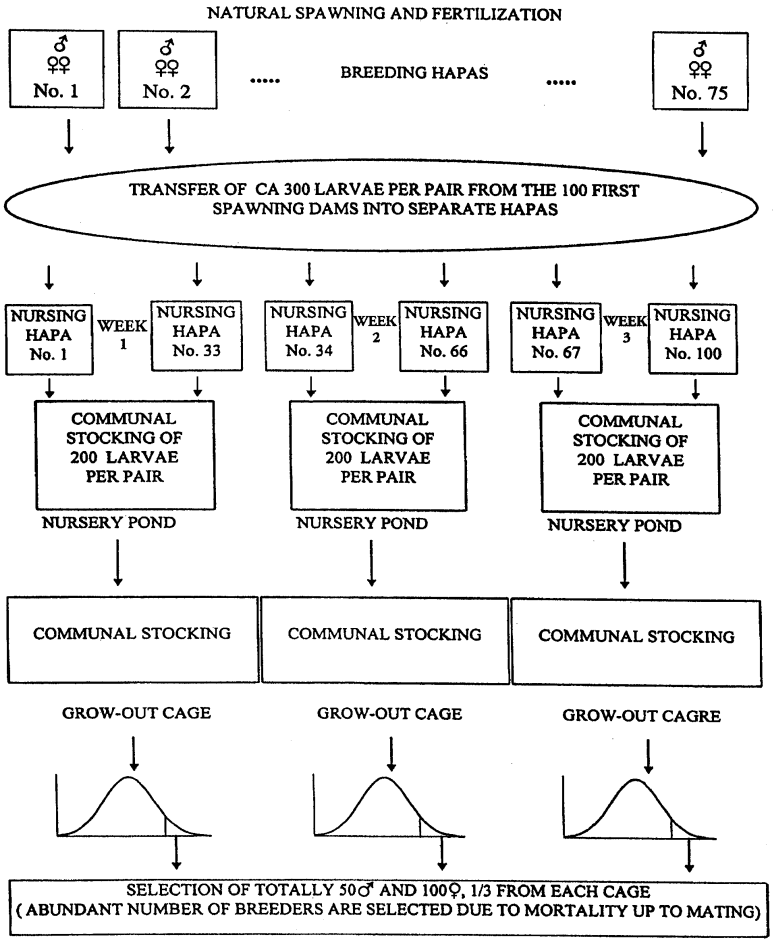

NATURAL SPAWNING AND FERTILIZATION
♂ ♀♀ No. 1
♂ ♀♀ No. 2
.....
BREEDING HAPAS
.....
♂ ♀♀ No. 75
TRANSFER OF CA 300 LARVAE PER PAIR FROM THE 100 FIRST SPAWNING DAMS INTO SEPARATE HAPAS
NURSING HAPA No. 1
WEEK 1
NURSING HAPA No. 33
NURSING HAPA No. 34
WEEK 2
NURSING HAPA No. 66
NURSING HAPA No. 67
WEEK 3
NURSING HAPA No. 100
COMMUNAL STOCKING OF 200 LARVAE PER PAIR
NURSERY POND
COMMUNAL STOCKING OF 200 LARVAE PER PAIR
NURSERY POND
COMMUNAL STOCKING OF 200 LARVAE PER PAIR
NURSERY POND
COMMUNAL STOCKING
COMMUNAL STOCKING
COMMUNAL STOCKING
GROW-OUT CAGE
GROW-OUT CAGE
GROW-OUT CAGRE
SELECTION OF TOTALLY 50♂ AND 100♀, 1/3 FROM EACH CAGE
(ABUNDANT NUMBER OF BREEDERS ARE SELECTED DUE TO MORTALITY UP TO MATING)

Figure 16.4 Mating of the parent broodstock and rearing of progeny until selection. Reproduced from Gjedrem et al. (1997) by permission of WorldFish Center.

The main requirement of this strategy is individual tagging of all breeding candidates. Identification of the animals is necessary for keeping pedigree records and is a prerequisite when family testing is applied. When individual selection is practised it is possible to have efficient breeding system without tagging the animals but it will require

standardization of number in each family as described by Gjerde et al. (1996) and as illustrated in Figure 16.4 and 16.5. A breeding program without identification of animals and no standardization of progeny groups will usually result in a rapid build up of inbreeding.

If pedigree records are kept, it is then possible to use combined selection strategies, utilising the performance of relatives to determine individual breeding values. The large number of half and full-sibs that may be produced simultaneously in fish and shellfish breeding programs, will increase the accuracy of estimates of individual breeding value. The use of sib records is also of great importance when selecting for traits that may not be recorded in the breeding candidates like quality traits, or traits that only may be quantified in frequencies such as survival rate, disease resistance and age at sexual maturation.

Systematic environmental effects like differences in age and between cages or ponds should be kept as low as possible. The records should be adjusted for these systematic effects in order to reduce them, see Chapter 5.

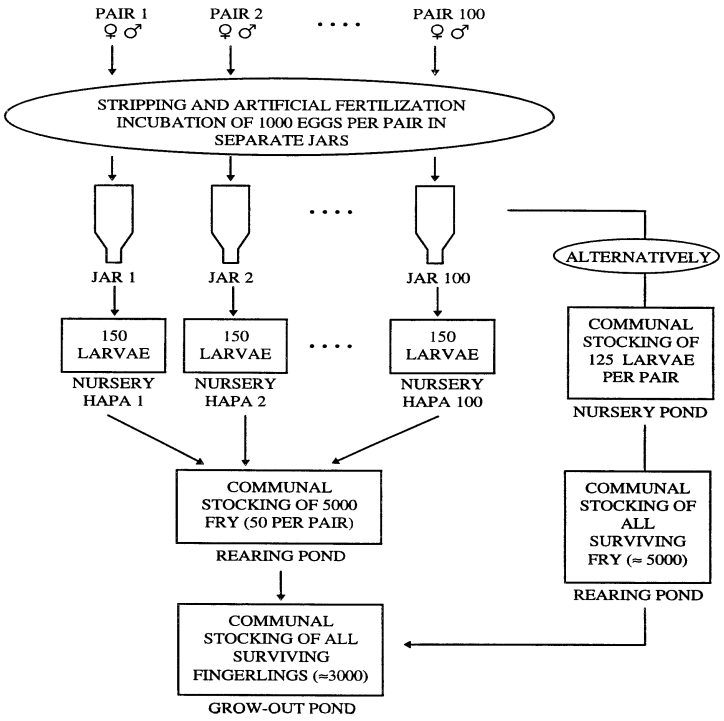


Figure 16.5 Mating and reproduction of the parent broodstock and rearing and testing of the progeny. Reproduced from Bentsen et al. (1996a) by permission of WorldFish Center.

16.13 Registration of records

Records, measures and scores, of the breeding goal traits should be taken at the time described in this goal. It is essential to work with high accuracy in all phases of recording in order to avoid errors, both systematic and random. All personnel involved should be trained in order to reduce errors in the recording procedures. Care must be taken not to make errors when transferring data into computers or other equipments. During the time of recording the animals should be anaesthetized to avoid stress and recording errors. Incorrect records will of course give false values of breeding candidates and reduce genetic gain.

Growth rate should be recorded as body weight at time of marketing or as carcass weight. As a selection criterion we are usually not interested in growth rate during certain periods of life, which may be estimated as specific growth rate. The growth from hatching to market size is the trait of economic importance. Therefore age differences among genetic groups should be minimised in order to make fair comparisons. If it is not possible to produce all animals with the same age, body weight should be adjusted for age differences between family groups. The problem with age differences may partly be overcome by keeping age groups separate during grow out, as illustrated in Figure 16.4. By assuming equal genetic merit in the different age groups, adjustment for environmental differences may be obtained by using deviation from age groups averages. Another possibility is to select equal number of broodstock from each age group.

In on farm trials there may be large variation between farms and even between cages or ponds within farms in different traits. In order to analyse such data statistically they may be adjusted by taking deviations from farm and or cage averages. As a first step of the analysis, the data should be tested for unreasonable outliers, and the steps should be taken to remove such errors.

Feed conversion efficiency has not been included in selection programmes with fish because of problems with recording feed intake of individuals or full-sib family groups (Gjedrem, 1983a). Short-term feed intake in individual fish can now be recorded by several methods, like video recording (Brännäs and Alanärä, 1992), “labelled feed” techniques using radioactive isotopes (Storebakken et al., 1981), X-radiographic method (Talbot and Higgins, 1983; Jobling et al., 1995) or coloured feed (Johnston et al., 1994). However, it is necessary to record the feed intake during several weeks or months when studying the feed conversion efficiency in fish. Long-term feed intake can only be recorded accurately for groups of fish reared in separate tanks or cages by over-feeding and collection of uneaten feed (Helland et al., 1996). Since the rearing environment will change in different tanks or cages, each full-sib family should be tested in a minimum of two tanks or cages to reduce the environmental effects common to full-sibs. As a consequence, it will be very resource demanding to record and select directly for improved feed efficiency in fish.

Since survival rate under rearing conditions usually show very low heritability, Gjedrem (1995b) proposed to apply challenge tests using standardised methods for the most serious diseases. For this purposes aquatic species with their high fecundity have a great advantage compared with livestock. For each disease a sample of animals from each full-sib family should be drawn for separate challenge tests. The test animals should be tagged before all families are mixed into one container where the pathogen in question is added. During the testing period dead animals should be recorded including time of death. The measure of resistance against a particular disease should be percent mortality of each family when 50 % mortality has been reached in the test. An alternative measure could be average time to death for each family when the challenge test lasted as long as fish were dying. The results from the challenge test should be used as criterion for estimation of breeding values for the families involved.

A challenge test with introduction of pathogen in the water for example by using cohabitants, should be preferred to introduction of the pathogen through feed or through injection. By introducing the pathogen in the water all defence mechanisms of the fish will be tested including mucus and skin. Relatively small animals are usually used in challenge tests because they are cheap, need little space, and can easily be handled. The results obtained may therefore not be representative for larger animals and animals reared in different environment. Therefore the success of selection under farming conditions will depend on the genetic correlation between survival after challenge test and resistance to the disease under commercial farming condition. As an example Gjøen (1996) report a high genetic correlation (r_G = 0.95) between challenge test against furunculous of 20-40 g fingerlings of Atlantic salmon in fresh water and a field test after 2½ months in marine cages. A challenge test for sea lice (*Lepeopheirus salmonis)* has been developed by AKVAFORSK (Kolstad et al., 2004). The genetic correlation between challenge test and natural attach in sea cage was high ($r_G = 0.80 \pm 0.21$).

Survivors from families with highest survival rate in challenge test would be interesting candidates for selection of broodstock. However, since they may be carrier of the disease, they are usually not allowed used for breeding. Uninfected siblings from the top resistance families kept at the breeding station should therefore be selected as broodstock. However, in the chase of sea lice it is possible to use the fish with the lowest number of sea lice during challenge test as broodstock.

Fat percentage can be measured by chemical methods and also by Computerised tomography. This technique can also be used to measure fat distribution (Gjerde, 1987; Rye, 1991). An alternative method for measuring fat % is to use the Torry fat meter (Kent, 1990), which is inexpensive and portable, but does not have the same accuracy as computerised tomography. AKVAFORSK has developed an image analysis which showed high accuracy for rainbow trout, R^2 = 0.83 between image and chemical analysis of fat in filet (Rørvik personal communication).

Flesh colour is of importance in salmonid and some shellfish. The retention rate of dietary carotenoid is low but show genetic variation, Table 5.7. Subjective scores may be used to judge flesh colour. However, it has been found that subjective judgement is

an unreliable method of recording colour. Gjerde and Gjedrem (1984) estimated heritability for colour score to be 0.01 and 0.06 in Atlantic salmon and rainbow trout, respectively. Applying an exact technical device by using a colormeter reading (Minolta Chroma Meter CR-300) the accuracy was much improved with a heritability of 0.47 (Rye et al., 1994). The image analysis showed a high degree of accuracy in rainbow trout filets, R2 = 0.86 between image and chemical analysis of astaxantine (Rørvik personal communication).

Texture of flesh is a quality trait of importance both in fresh as well as in refined products. Shahidi and Botta (1994) describe texture as firmness, cohesiveness and elasticity and is measured by different instruments using device for mechanical pressure. One of the instruments are Texture Analyser, TA-XT2 (Mørkøre and Rørvik, 2001). No instruments have been developed to measure taste of fish and shellfish. The only possibility is to do organoleptic judging which requires trained personnel and is very costly. Lack of methodology to record traits on live breeding candidates represents a limitation in selective breeding in aquaculture (Gjerde and Rye, 1997). It is therefore of importance to develop methods for making individual selection possible also for other traits.

16.14 Systems for selection procedures

16.14.1 INDIVIDUAL SELECTION

As described in chapter 11 it is essential to use a large number of broodstock in each generation in order to avoid inbreeding when individual selection is applied. Further it is also important to standardise the number of progenies from each breeding pair otherwise some pairs may have few progeny, while others may have many.

16.14.1.1 Example of natural mating, Figure 16.4
The mating of selected broodstock, one male and two females, may take place in hapas or tanks. Larvae from each full-sib group are transferred from breeding hapa to a nursing hapas and standardised, for example to 300 larvae each. During the rearing period in the nursing hapa, mortality may be considerable. As mortality declines the numbers are standardised again before the larvae are stocked into grow-out ponds or cages for communal rearing. As indicated in Figure 16.4 one can reduce age differences by stocking the families into several ponds or cages according to age.

16.14.1.2 Example of artificial fertilisation, Figure 16.5
Equal numbers of fertilised eggs from each pair are put into hatching jars or hatching trays. After hatching a sample of larvae are transferred into a nursery hapa or tank for startfeeding, usually a period with high mortality. When mortality is reduced a constant number from each family is transferred into ponds or cages for communal rearing until market size. An alternative and more simple method is illustrated on the right side in Figure 16.5 where larvae are stocked directly into communal ponds or cages, in situations with low mortality during startfeeding period this method could be used.

The aim of the procedures described above is to end up with the same number of progeny from each family in the grow-out ponds or cages, with minimum differences in age.

16.14.2 FAMILY SELECTION

16.14.2.1 Example of a nested mating design

Production of 50 half and 100 full-sib families is shown in Figure 16.6. In carp and other species hormone treatment should be used to reduce age differences between the groups. From each spawn at least 1000 eggs should be transferred into a hatching jar. Newly hatched larvae from each family should be transferred into one hapa or tank and reared there until they reach tagging size, 3-15 g depending on method of tagging and species. As an example 160 fingerlings from each family can be tagged. After tagging all families should be communally stocked in a pond or cage in order to reduce environmental differences between families as much as possible.

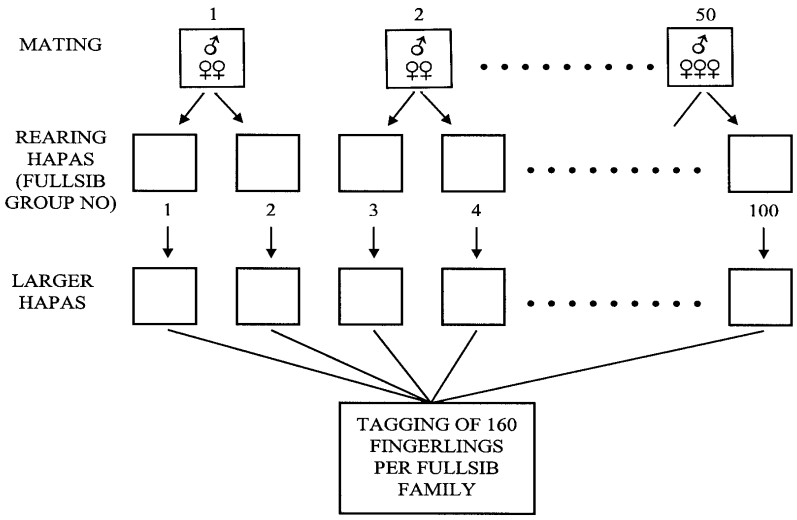


Figure 16.6 Mating and rearing design for one testing round of family selection. Reproduced from Gjøen et al. (1997) by permission of WorldFish Center.

16.14.2.2 Family testing in different environments and recording of several traits

Figure 16.7 illustrates testing of families at different test stations and for different traits. The purpose of these test stations is to study the ranking of families in different environments. From each family 100 fingerlings are communally reared in grow-out ponds or cages at the breeding centre. These animals can be considered as potential broodstock. In order to test the families in different environments and under practical farming conditions, samples of 15 fish from each family are reared in cage culture and 15 fish from each family are reared in pond culture. The fourth sample of fish from each family are proposed to be used for challenge tests, to study resistance against a disease. The other samples may be used to measure other traits of importance.

During the testing period records are taken as described in the breeding goal. In the example presented in Figure 16.7 body weight, age at sexual maturation and disease resistance are tested and recorded. All records are transferred to the breeding centre for estimation of breeding values. If the records at the test station are taken earlier than at the breeding station, these records can be used to make a preliminary ranking list of breeding values for the families. This ranking list may later be used for preselection of broodstock at the breeding centre.

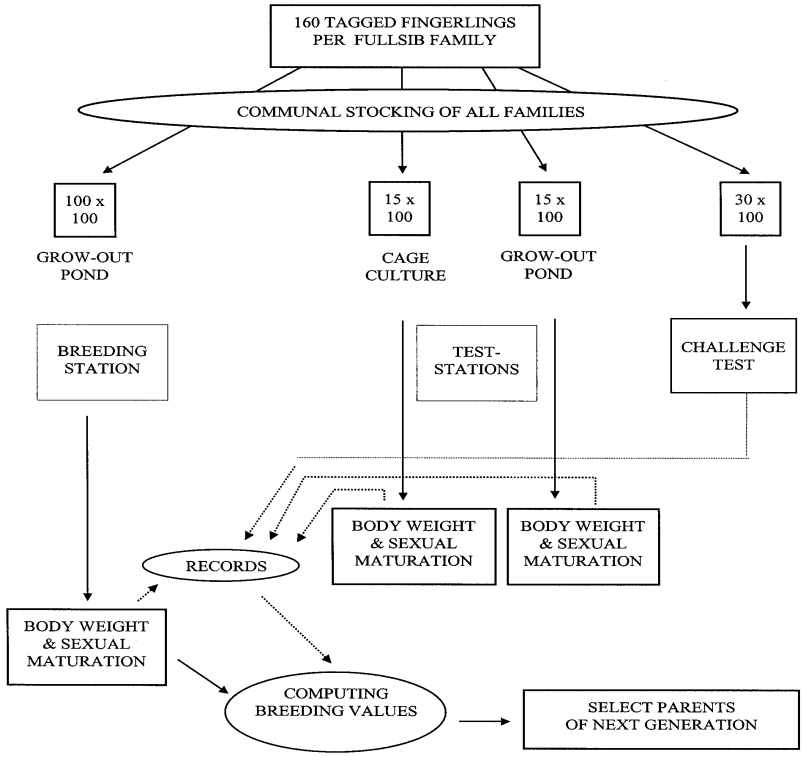


Figure 16.7 Testing, recording and computing of breeding values. Reproduced from Gjøen et al. (1997) by permission of WorldFiah Center.

16.14.3 PROGENY TESTING

For multi spawners progeny testing can be of interest particularly since both sexes can be tested. Selection of broodstock can not take place before the traits has been recorded on their progeny which means that the generation interval will be increased. The advantage of progeny testing is that the accuracy of breeding value can be very high, approaching 100% as number of progeny becomes very large.

16.15 Estimation of breeding value

Recording of data in a breeding program may be laborious and time consuming because number of animals will be large and because several traits may be included in the breeding goal. Therefore a computer with large capacity is needed to store and process the large mass of data recorded. It is essential that recording and input of data is well organized and planned. Calculation of breeding values should use all available data to get the best possible estimate of genotype for the individual animal or family. The

accuracy of estimated breeding values will vary between 0 (No genetic variation present) to 100 %.

The accuracy of breeding values depends partly on the quality of data recorded. Therefore a detailed description of time for recording, how to measure the traits, how the records should be stored without errors, and how the identification of the animals should be read and stored is necessary. As all information is installed in the computer, the data must be screened for possible errors (unreasonable figures) and adjusted for systematic environmental effects.

In Chapter 13 the procedures for estimation of breeding values is outlined and discussed.

16.16 Selection of broodstock

Selection of broodstock is an important part of a breeding program and requires detailed planning. The simplest situation is when only individual selection for growth rate is practised and when tagging is not used. If sex differences are small, a grading system can be used to select the heaviest animals as broodstock. The selection should take place at market size and the selected animals should be reared till maturation. Calculating number of selected broodstock should take into account capacity of the breeding station, demand from the market and death rate from selection to maturation. If there are considerable size differences between sexes when selection take place, and type of sex is not visible, one has to select a large enough number of animals to be ensure a sufficient number of the smallest sex are selected.

16.16.1 An example of individual selection, Figure 16.4
The animals in the three cages may be considered as genetically equal. If 150 females and 75 males are needed as broodstock, one may select the heaviest 50 females and the heaviest 25 males from each cage. Mating males from one cage to females from another cage, mating of full-sibs can be avoided.

16.16.2 An example of family selection, Figure 16.7
It is proposed to first take records at the test stations and then use these data to calculate family indexes, and thus get a ranking of all families under test. These indexes may be used when records are taken at the breeding station and for selection of potential broodstock from the best families. For example males heavier than 2σ above the average from the 10 to 15 best families, and females heavier than 1σ above the average from the 15 to 20 best families could be selected. When all data from the breeding station is recorded and added to data from test stations, breeding values for individual animals can be calculated and used as basis for the final selection of broodstock.

16.17 Controlling the amount of genetic gain

Establishing a procedure for the control of genetic gain in a breeding program is not required in order to obtain selection response. However, including a control mechanism in a breeding program makes it possible to continuously check if genetic change is according to expectation and if not, make necessary adjustments of the program. A breeding program will always have limited testing capacity and therefore one must

decide how much should be used for control groups since this will reduce number of genetic groups tested.

Methods for measuring genetic change have been discussed in Chapter 15. The most commonly used methods are:

- Unselected control group
- Divergent selection
- Repeated mating
- Use of average breeders
- Genetic trend analyses
- Response based on selection paths

16.18 Constraints to breeding programs

16.18.1 NEGATIVE CORRELATED EFFECTS

When several traits are included in the breeding goal, genetic correlations between the traits will be part of the equations for selection indexes. When there are genetic correlations between traits in the breeding goal and other traits not included in the goal, selection will change these unrecorded traits as well. In some cases this correlated response may be advantageous, but it could also be detrimental to the production system. In farm animals Rauw et al. (1998) have given several examples of negative side effects of selection such as leg problems in poultry and pigs and reproductive traits in dairy cattle, turkey and poultry. It is therefore important that undesirable correlated responses are carefully studied. In aquatic species they are most likely to appear in fitness traits like quality of gonads and vitality in the early life stages. When negative effects are detected the breeding program must be changed, and action should be taken to repair the damage already done.

16.18.2 CHANGE OF BREEDING GOAL

The importance of growth rate and disease resistance in the breeding goal is unlikely to change over time. However, for traits like meat quality, it might happen that the consumers change their opinion and start to prefer a quality different from the one selected for. This type of risk is difficult to avoid. However, it is very important for a breeding company/organization to study carefully trends in market preferences particularly concerning product quality.

16.18.3 SPREAD OF DISEASES

When a breeding program is centralized as discussed in chapter 18, there is a risk that infectious diseases can be transferred from the breeding centers to the industry. To avoid this it is extremely important that the breeding centres and multiplier stations work at a high hygienic standard and carefully control import of new fish, following strict

quarantine regulations. Everything possible must be done to produce healthy animals for the industry.

It has been argued that selecting for disease resistance can produce carriers of diseases and thus be a source of infection. It is, however, most likely that bacteria and virus are dependent on susceptible hosts.

16.18.4 GENOTYPE-ENVIRONMENT INTERACTION

In section 14 it was concluded that it is not possible to draw a general conclusion about presence of genotype-environment interaction and therefore it should be studied at the start of a breeding program. If the interaction is considerable it may be necessary to develop more than one strain. However, there is the possibility that strains over time may become more sensitive to environmental variation. To avoid such surprises one should periodically investigate possible genotype-environment interactions. The risk of genotype- environment interaction detrimentally affecting the breeding program can be reduced by always testing breeding values under common farming conditions and on several private farms.

16.18.5 INBREEDING ACCUMULATION

When reproducing species with high fecundity and in a closed population, there is always the possibility of rapid accumulation of inbreeding. By using tagging of animals and by keeping pedigree records, it is easy to avoid mating of close relatives and thus secure a slow build up of inbreeding. When individual selection is used without identification of animals, it may be difficult to avoid inbreeding. However, it is possible to reduce the inbreeding problem by dividing the breeding population into at least two subpopulations. Broodstock should be selected and produced within each subpopulation. Mass production of progeny can then proceed by crossing subpopulations. Introduction of unrelated stocks is another efficient way of reducing inbreeding. Inbreeding is discussed further in Chapter 6.

16.18.6 REDUCTION IN GENETIC VARIATION

Another constraint mentioned in connection with selection programs is a possible reduction of genetic variation in traits under selection. According to Falconer (1960) inbreeding will reduce the genetic variance. This is another good argument for the avoidance of inbreeding. Several authors have discussed the effect of selection on genetic variation. Fimland (1979) showed that a small reduction in the genetic variation is expected during the first generations of selection, after which a stable level is reached. The level at which the genetic variance stabilises depends on selection intensity of parents and the accuracy of selection. If the accuracy of selection does not exceed 0.6 the stabilized level will not be lower than 80 %. The true variation will be re-established at the initial level if is terminated. In relatively large populations, Enfield (1979) selected for pupa weight in tribolium in 120 generations and found no reduction in genetic- and phenotypic variation during the selection period. Long term breeding

programs in large populations of farm animals do not normally result in reduction of genetic variation.

16.19 Dissemination of genetic improvement

In order to maximise the benefit from the breeding program, the genetic improvement must reach the farmers without delay. The product for dissemination will usually be improved semen, eggs, fry or broodstock. There are mainly two routes for dissemination:

1. One route is illustrated in Figure 16.8. Semen, eggs, fry or broodstock is transferred directly from the breeding station to the farmers. This can be surplus eggs or fry from the selected broodstock, spent breeders or next best breeders
2. A second route is to establish multiplier stations for production of semen, eggs or fry from broodstock originally produced at the breeding station. The multiplier stations should produce the necessary number of broodstock and select for growth rate, Figure 18.1

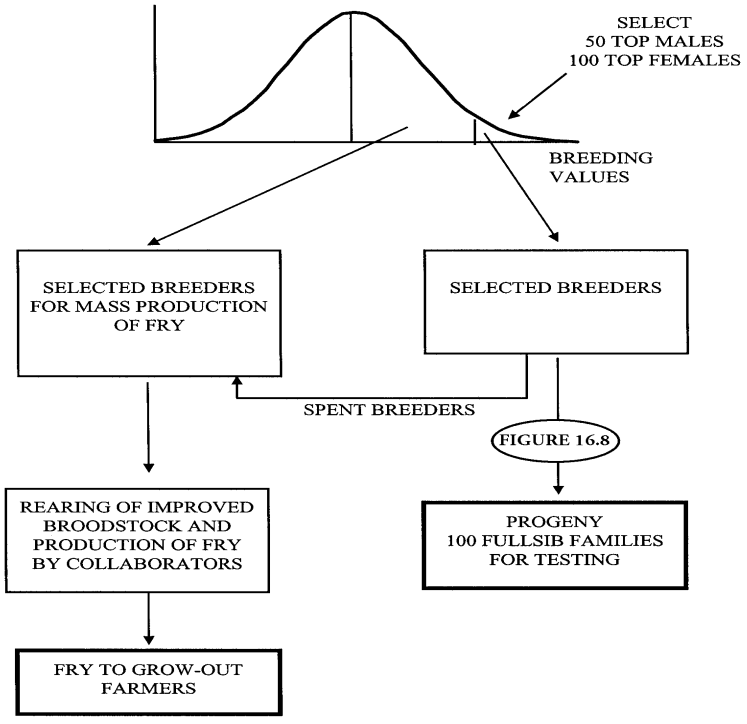


Figure 16.8 Selection of breeders and dissemination of gentically improved stocks to farmers, in each round of testing. Reproduced from Bentsen et al. (1996b) by permission of WorldFish Center.

The breeding work must be well organised in a breeding cooperative or a breeding company. Capital is needed to work out a breeding plan and establish necessary testing facilities. One should take into account that it takes time to improve the stock to a degree where the benefit is obvious to the farmers. Further it takes time to scale up the production of improved animals and establish a market for the product.

16.20 Combining selective breeding and marker assisted selection

Marker assisted selection has been discussed in Chapter 19. It is of particular interest for traits with low heritability and for traits which can not be measured on live animals like quality traits. Another advantage of marker assisted selection is that it may be possible to identify breeding candidates of high genetic value at an early age.

16.21 Summary of main elements in a breeding plan

A breeding plan is never a final document and should be continuously revised and improved as new knowledge is obtained through experience and research.

The main elements in a complete breeding plan are:

- Define the breeding goal
- Form a base population with broad genetic variation
- Select breeding strategies
- Chose selection methods
- Identification of animals
- Select mating design
- Protocol for testing strategies
- Protocol for recording of all traits in the breeding goal
- Describe systems for selection procedures
- Estimate breeding values
- Protocol for selection of broodstock
- Protocol for control populations
- Avoid constraints and negative effects of selection
- Plan for dissemination of improved stock and organization of breeding work

17 ORGANISING BREEDING PROGRAMS

TRYGVE GJEDREM AND TERJE REFSTIE

17.1 Practise in farm animals

A modern breeding plan is quite complicated and usually deals with a large population of animals. To be efficient the work of breeding and distribution of the breeding product (genetically improved animals) must be well organised. In many countries the farmers have established breeding cooperatives to carry out their breeding programs. On behalf of their members the cooperative builds the necessary test and breeding stations and develops recording systems in order to get information/records as the basis for estimation of breeding values. In other countries this is organised by private companies

In dairy cattle for examples performance testing for meat production traits of young bulls take place at special test stations. After selection the bulls are progeny tested for milk yield of their daughters, which are born and reared on private farms. The basis for estimating breeding values of bulls are the records of milk yield and milking ability of their daughters and on their own growth performance. Use of AI (Artificial insemination) and deep-freezing of semen has made it possible to develop very efficient breeding programs in dairy cattle. Since it is costly to run a breeding program in cattle with progeny testing of bulls, it is essential that each program cover a large population of cows. There are also large private companies in dairy cattle breeding, which have established breeding programs on a completely commercial basis and frequently cover several countries and work in competition with the breeding cooperatives. The breeding product in cattle is sperm from highly selected bulls, stored and transported deep-frozen.

In pigs private companies are more and more dominating the market for breeding material. These companies give farmers cooperatives, which are similar to those in dairy cattle, hard competition. To day a few large companies cover a substantial part of the world market. The breeding product in pigs is semen and or young breeding boars and gilts.

In poultry a few large companies control the whole world market of both egg-laying hens as well as for broilers. The breeding product in poultry is usually fertilised eggs from highly selected animals from different lines, and the farmers use hybrid animals.

Breeding programs for sheep vary considerably. Individual selection combined with crossbreeding of breeds and lines has dominated. Improvement programs have usually focused on growth rate and fertility. A major problem in sheep breeding has been low conception rate from use of AI. In Norway a breeding program using natural mating has been applied since the mid 1960's with satisfactory results (Steine, 1980).

For aquatic animals with high fecundity breeding programs will include breeding stations (breeding nucleus) where the mating of selected animals take place, test stations

T. Gjedrem (ed.), Selection and Breeding Programs in Aquaculture, 279–285.
© 2005 *Springer. Printed in the Netherlands.*

for recording production traits under ordinary farming conditions, and multipliers for dissemination of genetically improved animals. The organisation of the breeding program will very much depend on what selection methods are used, the fecundity of species and the total production in the industry utilizing the improved stock.

The basic element in breeding programs for farm animals is selection in order to improve economically important traits. Selection methods most commonly used are individual selection, family selection and progeny testing of sires. In chickens, and to some extent also in pigs, crossing of genetic lines is practised to exploit non-additive genetic variance. When breeding programs include crossbreeding, the breeding company will keep animals from their lines or strains, and sell crossbred animals. In this way the breeding companies protect their inbred lines, as the farmers must buy crossbred animals from the breeding company. The crossbreds, which the farmers buy, can not be duplicated.

17.2 Organisation of breeding programs in aquaculture species

17.2.1 ATLANTIC SALMON IN NORWAY

In 1971 AKVAFORSK started genetic research with Atlantic salmon, and the first selection of broodstock was performed in the fall of 1975 at the research station at Sunndalsøra. Four synthetic base populations were formed based on 40 wild Norwegian river strains. The four populations or year-classes were made because Atlantic salmon has a generation interval of four years. The two first generations of selection were for increased growth rate only, while selection against early sexual maturation was included from 1982, disease resistance from 1991 and flesh quality traits from 1994. Combined individual and family selection was performed testing about 200 families each year.

Table 17.1. Year-classes/populations of Atlantic salmon in the Norwegian breeding program.

Generation	1	2	3	4
0	1972[1)]	1973	1974	1975
1	1976[2)]	1977	1978	1979
2	1980	1981	1982[3)]	1983
3	1984	1985	1986[4)]	1987
4	1988	1989	1990	1991[5)]
5	1992[6)]	1993	1994[7)]	1995
6	1996	1997	1998	1999
7	2000	2001	2002	2003

1) Year for hatching
2) Selection for growth rate started
3) Selection against early maturation started
4) Second breeding station established
5) Selection for disease resistance started
6) AquaGen AS was established
7) Selection for flesh quality started

AKVAFORSK continued the selection program at Sunndalsøra for 2½ generations and started to disseminate eyed eggs and smolts from selected parents. Then the Fish Farmers Sales Organisation (FOS) and Norwegian Fish Farmers Association (NFF) were invited to take the responsibility for the breeding program. In 1985 FOS and NFF built facility for a second breeding station at Kyrksæterøra (NFA) and AKVAFORSK transferred selected genetic material from their own breeding station at Sunndalsøra to the new breeding station. AKVAFORSK continued to be a partner in the breeding program until 1992. At that time AquaGen AS was established as a limited company. In 2000 a second breeding company was established, SalmoBreed AS, by the industry.

The breeding stations work as the breeding nucleus where family testing and selection takes place. The breeding station is the key for genetic improvement, which takes place every generation in each of the four year-classes, Table 17.1.

To ensure large enough production of eyed eggs to the industry, multiplier stations were established. The multipliers are all private fish farms with hatcheries for the production of eyed eggs and production of broodstock using genetically improved smolts from the breeding company. A contract between the breeding company and the multipliers regulates the cooperation. The number of smolts delivered from the breeding company

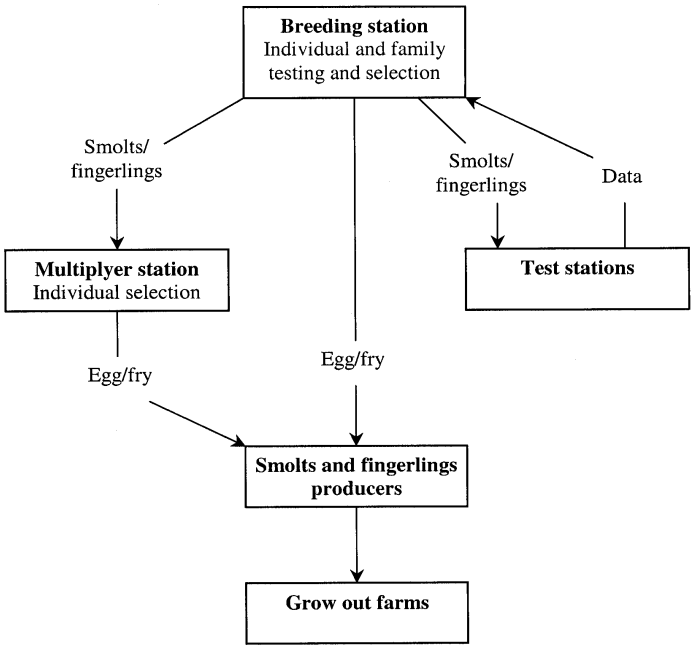


Fig. 17.1. Dissemination of genetic gain from Norwegian salmon breeding program.

to the multipliers is sufficient to allow them to practise individual selection for growth rate. The multipliers together with the breeding stations produce eyed eggs, which are sold to hatcheries for smolt production and for later grow-out in sea cages. The price for eyed eggs includes royalties to the breeding company to cover the expense for running the breeding program.

Dissemination of the genetic gain from the two breeding stations to fish farmers is shown in Figure 17.1.

17.3 Structure of breeding programs.

Although the basic breeding theory is the same for all animals the breeding programs vary considerably. There are several reasons for this variation:

- fecundity of the species
- different breeding goals
- whether traits are sex limited or not
- economic value of each animal
- whether fertilisation take place in uterus or outside the female
- cost of broodstock production

17.3.1 FECUNDITY OF THE SPECIES

Fish and shellfish species are extremely fecund with production of thousand and even millions of eggs per spawn as compared with cattle, which usually release one egg per oestrus.

In dairy cattle where breeding programs usually are quite efficient, most of the cows are included in the breeding population because records of milk yields are used to progeny test the bulls. A relatively large number of daughters are needed to get reliable estimates of breeding values because dairy cattle herds are small in Norway and with large environmental differences. To overcome this problem the daughters milk yield is taken as deviation from the heard average. To evaluate meat-producing traits it is usual to establish performance test stations for young bulls to study growth rate, feed conversion efficiency and meat quality traits. Such test stations are quite expensive to run.

As is explained above testing of breeding values and selection for fish and shellfish can take place in one breeding station, which can serve a large population. For example in Atlantic salmon in Norway there are two breeding stations and they supply, together with the multipliers, most of the industry with improved eyed eggs for a production of 500 000 tons of salmon and rainbow trout. However, the capacity of this breeding program is many times larger than the present production. In addition some commercial farms are used as test stations where production traits are recorded. This ensures that selection is based on production records taken under normal commercial farming conditions.

17.3.2 DIFFERENT BREEDING GOALS

There may be considerable difference in breeding goals among species and even within species. If one or more traits in the breeding goal can not be recorded on individual animals alive, family selection or progeny testing must be applied. The structure of the family breeding or progeny testing program will be much more complicated compared with a breeding program utilising individual selection only.

17.3.3 THE CHASE OF SEX LIMITED TRAITS

For species with sex limited traits of economic importance, such as milk yield in cattle, progeny testing must be utilized to estimate breeding values of males. The downside of progeny testing is that the generation interval is lengthened which dramatically reduces the genetic gain per generation. Progeny testing of males is particularly applied in cattle, sheep and goats. In many salmonid species the fish do not survive reproduction and therefore progeny testing can not be used unless cryopreservation of gonads are included as part of the breeding program.

17.3.4 ECONOMIC VALUE OF EACH ANIMAL

In farm animals, particularly for the largest species, the value of each animal is high. This restricts the flexibility of a breeding program. For example it is not possible to test for disease resistance by applying challenge tests where many animals will die, for cattle, pigs, sheep and goats, while it could be possible for chicken. In fish and shellfish the economic value of each animal is low so challenge tests for different diseases is possible. This opens the possibility for developing disease resistance strains. The same argument will to some extent also apply for some carcass traits.

17.3.5 WHETHER FERTILIZATION TAKES PLACE INTERNALLY OR EXTERNALLY THE FEMALE

The cultured fish and shellfish species all have external fertilisation. This makes it possible to practise a variety of mating systems as described in Chapter 12. The eggs from one female can be fertilised by sperm from many males and the milt from one male can fertilise eggs from many females. With this flexible mating system it is convenient to study magnitude of different genetic components, additive, dominance, epistatic effects, maternal effects together with genotype-environment interactions.

17.3.6 COST OF BROODSTOCK PRODUCTION

Selection for high growth rate in young animals will usually result in a correlated response in mature body weight. In farm animals this will have an adverse effect because the cost of maintaining breeders will increase. The cost of growing and keeping breeders is substantial in many species. Large (1976) has estimated the percentage of total food which is eaten by female breeders to be 72 % in meat sheep, 52 % in beef cattle, 33 % in pigs and 10 % in poultry. The figure for fish range from less than 1 % for species

slaughtered close to maturation to about 5 % (Kinghorn, 1983a). This very low cost of reproduction indicates that breeding programs for aquatic species will be relatively less expensive compared with farm animals. This should stimulate fish farmers to invest in breeding programs. For fish and shellfish it is uncomplicated to increase growth rate because broodstock are generally kept for a short period of time until they reach maturity.

17.4 Farmer cooperatives - breeding companies

In most countries animal husbandry has traditionally been small scale farming. And historically the small farmers cooperated to keep males, as one male could serve several herds. As modern breeding theory become known the farmers formalised their cooperation and formed larger groups and developed breeding programs. Sometimes they got support from their government but usually they took responsibility and financed the necessary investments and covered the running costs. The breeding plans were usually developed with assistance from researchers with knowledge in breeding theory. As the breeding cooperatives and the researchers gained experience and new technology was developed, they were able to make more efficient breeding programs. Later modern data technology were used to store the production records and the large number of collected records could be combined and utilized much more efficiently in estimation of breeding values.

Farmers cooperative breeding programs have proven to be a good model for the organisation of breeding programs. This model is common in dairy cattle where recording of production and testing of breeding values are decentralised and takes place on the farms. In species where recording and testing of breeding values is centralised and takes place in breeding stations, possibly combined with test stations, it is more common that private companies have their own breeding program. This is particular true in poultry where a few breeding companies cover the total world market with improved animals.

In fish and shellfish breeding programs could be organized both as private companies and either as cooperatives where farmers are well organised. However, when the breeding work is done by companies, farmers and farmers organisations should be shareholders in order to influence the direction of the breeding work and breeding policy.

17.5 Initiative to start breeding programs

Gjedrem (1997b) estimated that only about one percent of fish and shellfish produced in 1994 originated from efficient breeding programs. Taking into account the great economic advantage together with much better utilisation of resources of feed, land and water when genetic improved animals are used, it is urgent that breeding programs are started without delay in all countries and for all aquacultural species. It has been shown

that the benefit relative to cost is very high for efficient breeding programs. According to Barlow (1983) and Mitchell et al. (1982) the cost/benefit ratio ranged from 1:5 to 1:50 in sheep, cattle and pigs. There is reason to believe that cost/benefit ratios will be still more favourable in fish and shellfish because of their high fecundity and the low cost of producing and keeping broodstock. With this in mind there should be no excuse for not starting breeding programs for any species in commercial production.

Fish and shellfish farmers should feel responsibility for starting breeding programs because they are the ones who will gain economically, by increasing productivity and by giving themselves a competitive advantage in relation to other products. Since it is difficult for individual farmers to start breeding programs, they should encourage the development of association in aquaculture to organise and start breeding programs. Initiative could also come from ministries, which have the responsibility to stimulate higher productivity in aquaculture production. In a number of countries researchers and research organisations have been the primary initiators of breeding programs.

17.6 National or local breeding programs?

There is no direct answer to this question. However, it is obvious that there is a great advantage for breeding programs to work with a large population of animals. The running cost will usually be reduced per unit as the breeding population increases. But there are also limitations to population size. International breeding programs will frequently meet national restrictions, which try to reduce introduction of diseases. For this reason it is foreseen that most breeding programs will work on a national basis.

It may also be necessary to divide a population within a country into subpopulations because the consumers ask for different products and product quality. However, since there always is considerable variation within a population, this request can frequently be met by grading the carcasses into quality groups before marketing. The problems usually encountered are to find exact and rapid methods for quality assessment.

It may also be appropriate to divide a population into subpopulations if there are considerable genotype-environment interactions. This means that ranking of families and other genetic groups are different in different environments or farming conditions. As discussed in Chapter 14 such interactions accounting for a considerable part of the variance are found in some investigations of common carp (Moav et al., 1975; Wohlfarth et al., 1983) while interaction for growth rate in salmonides, tilapia and shrimp seams to be of minor importance (Gunnes and Gjedrem, 1978, 1981; Eknath et al., 1993; Fjalestad et al., 1997). In conclusion it is not possible to predict the importance of genotype-environment interaction in advance and therefore this matter should be investigated in each instance.

18. CHROMOSOME ENGINEERING

TERJE REFSTIE AND TRYGVE GJEDREM

18.1 Introduction

In most fish and shellfish species fertilization takes place externally which makes it possible to manipulate number of chromosomes. Techniques have been developed to produce both haploids and polyploids with three sets of chromosomes (triploids) and even four sets of chromosomes (tetraploids). It is also possible to produce animals where chromosomes come from only the dam (gynogenesis) and from only the sire (androgenesis).

Gynogenesis is a natural form of reproduction in the teleost *Mollienesia formosa* (Hubbs and Hubbs, 1932), some population of crusian carp (*Carassius auratus gibelio*) and in several species from the family *Poeciliidae* in Mexico (Gold, 1979; Cherfas, 1981). The existence of distinct clones in laboratory fish has been demonstrated by tissue transplantation tests (Kallman, 1962). Perfect normal and fully viable broodstock of these species can be reared. In gynogenesis the development of the egg nucleus is triggered by the sperm of a closely related species which interacts with the egg but do not fuse with it. Thus the male nucleus makes no contribution to the developing embryo.

Spontaneous triploids have been identified in roach, *Rutilus rutilus* (Gold and Avis, 1976) and in hatchery populations of brook trout and rainbow trout (Allen and Stanly, 1978; Thorgaard and Gall, 1979).

The purpose of chromosome engineering is mainly to produce gynogenetic inbred lines, sterile triploids or polyploids. The work was originally started by Hertwig (1911) who inseminated frog eggs with spermatozoa inactivated by irradiation. Hertwig showed that eggs could develop without a male genetic input. This process has been called gynogenese (all-maternal inheritance), which is a special form of parthenogenese (development of an egg without fertilization). Surviving embryos were found to be mainly haploid with a low frequency of 1% being diploids with retention of the second polar body or blocking of first cleavage of the cell. It was later found that temperature shock, hydrostatic pressure or chemical treatments of inseminated eggs could increase the frequency of diploid embryos. A complementary process was later established by irradiation of egg chromosomes following normal fertilisation. The embryonic development involved then only chromosomes contributed by spermatozoa. This process has been called androgenesis (all-paternal inheritance).

By applying the same techniques but using normal sperm to fertilize the eggs, it is possible to produce polyploids. If the treatment is applied shortly after fertilization, triploids will be produced due to retention of the second polar body. If treatment is applied shortly before the first cell cleavage to suppress this process (event), tetraploids will be produced due to doubling of the chromosome number. Production of sterile

T. Gjedrem (ed.), Selection and Breeding Programs in Aquaculture, 287–299.
© 2005 *Springer. Printed in the Netherlands.*

triploids has been practised in some species to avoid developing early sexual maturation and tetraploids have been used to produce triploids by crossing normal diploids with tetraploids.

These methods were first applied to fish by researchers from the Soviet Union (Romashov et al., 1961; Golovinskaya, 1968), by Purdom in England (1969, 1970 and 1972a) and further expanded to include polyploidy. Chromosome manipulation has now been performed in many fish species and also in clam (Allen et al., 1982) and oysters (Stanly et al., 1984). Examples of review articles of the extensive research carried out in this field are: Purdom (1983) and Thorgaard (1992).

18.2 Gynogenesis

The purpose of producing gynogenetic fish is to develop highly inbred animals and a female stock. It is done by activation of an egg by genetically inert sperm. The sperm can be genetically inactivated by exposure to high doses of ionising radiation with around 100 000 rad or exposure of the sperm to ultraviolet light (Chourrrout, 1982a). The exposure of sperm to radiation or ultraviolet light will destroy and break the chromosomes in many pieces. However, the sperm is still mobile and has the ability to penetrate the egg and activate it to start the second phase of meiosis.

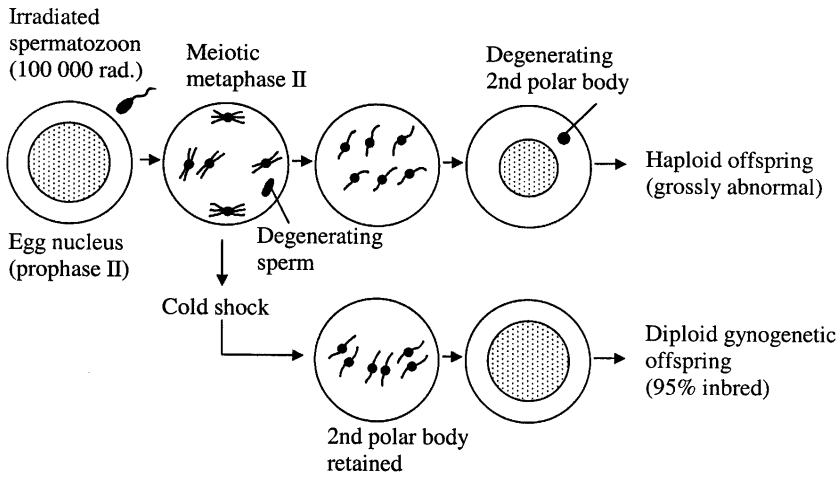


Figure 18.1. Diagrammatic representation of induced gynogenesis. Reproduced from Purdom (1993) by permission of Springer Science and Business Media.

It should be mentioned that in the normal cleavage of the nucleus, meiosis is halted after the first phase in which the first polar body is produced which later degenerates. The second phase of meiosis does not occur until after fertilisation when the second cleavage takes place, and again one of the two nuclei called the second polar body degenerates (Figure 18.1).

When the chromosomes in the sperm are destroyed, the activated egg contains only a haploid set of chromosomes after the completion of the second phase of meiosis. Some of these haploid eggs will develop into embryos, which frequently are abnormal and rarely survive until hatching. Haploid and diploid fish embryos are very different and can be easily separated which makes this system ideal for experimental purpose.

18.2.1 INACTIVATION OF SPERM

The first experiments to inactivate sperm used high dose of ionising radiation, which is necessary to inactivate the sperm but is somewhat hazardous and not readily available. An alternative was therefore an advance when Stanly (1976) and Chourrout (1982a) started to use ultraviolet radiation to inactivate the sperm cell. The use of ultraviolet radiation is now commonly used to produce gynogenetic embryos and the experience is that this gives better results than ionising radiation. An alternative way of producing gynogenetic embryos is to use sperm from a species unrelated to the experimental fish.

18.2.2 SHOCK TO PRODUCE GYNOGENETIC DIPLOIDS

Early work with amphibian (Frankenhauser, 1945), marine flatfish (Purdom, 1972b) and salmonids (Svardson, 1945) showed that it was possible to get a high frequency of diploids by applying cold shock. This was done by keeping newly fertilised eggs in a mixture of water and ice. However, Lincoln et al. (1974) did not succeed producing diploid gynogenetic larvae using cold shock in Atlantic salmon. Chourrout (1980) showed that heat shock was an effective alternative to cold shock in salmonids, a finding later confirmed by Thorgaard et al. (1981), Refstie (1983b) and Lincoln and Scott (1983). Hydrostatic pressure has also been used to produce diploid gynogenetic fish (Lou and Purdom, 1984b).

Timing of shock in order to produce diploids is important and must be worked out for each species. The shock is usually started 5-15 minutes after fertilization and should last for up to 20 minutes using heat shock or 20-120 minutes when cold shock is applied. There is a need for more accurate timing using heat shock or hydrostatic pressure than using cold shocks. For Atlantic salmon and rainbow trout Refstie (1983b) concluded that the highest frequency of gynogenetic fry were produced when Atlantic salmon and rainbow trout eggs were heat shocked at 24 and 26 ^{o}C, respectively, Figure 18.2. Lou and Purdom (1984b) produced a yield of up to 81% heterozygous diploid gynogenomes by applying hydrostatic pressure (55 000 kPa) for 10 minutes and 40 minutes after activation.

The antibiotic Cytochalasin B has been used to produce ploidy (Refstie et al., 1977; Allen and Stanly, 1979, 1981; Refstie, 1981). These authors found mosaics and the percentage of polyploid individuals fluctuated with concentration and treatment. Bolla and Refstie (1985) concluded that Cytocalasin B does not appear to be a useful chemical for the induction of polyploid fish. The percentage of polyploid individuals is too low and the level of triploids induced can be obtained in cheaper and more efficient way using heat shock or hydrostatic pressure.

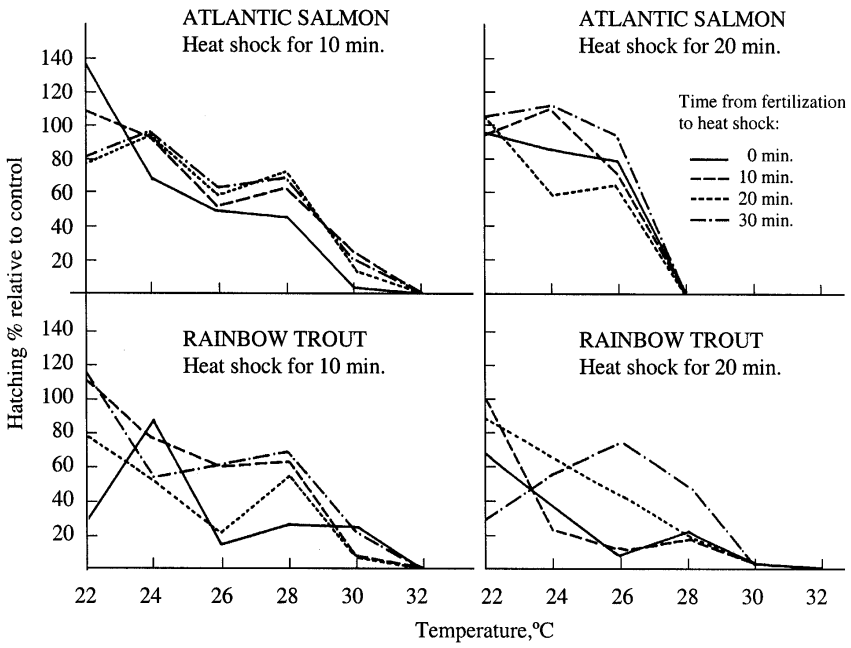


Figure 18.2 Hatchability relative to control for heat-shocked eggs of Atlantic salmon and rainbow trout at different temperatures and after time intervals from fertilization. Reproduced from Refstie (1983b) by permission of National Research Council, Canada.

18.2.3 PRODUCTION OF INBRED LINES

The diploid gynogenetic offspring contain only maternal genetic material. This is similar to self - fertilization, which produces progeny with an inbreeding coefficient of F = 0.50. The two sets of chromosomes in the diploid gynogenetic fish has the same origin and came from the duplication which occurred during the first phase of meiosis. As the result of crossing-over the inbreeding will not be complete and was estimated to F = 0.60 in one generation of gynogenese in common carp (Cherfas, 1981).

In species where females are the homogametic sex as in salmonids, all diploid gynogenetic fish will be females with only genetic information from its dam. In order to develop highly inbred lines it is essential to also be able to produce inbred males. Highly inbred males can be produced by hormone treatment of gynogenetic females. Good results have been obtained by adding the male hormone (17 α-methyl-testosterone) to the feed used to gynogenetic fry for the first two months on feed (Johnstone and Youngson, 1984). By this method genetic females have been sex-reversed into phenotypic males. The sex-reversed gynogenetic females will become sexual mature and produce viable sperm which fertilise eggs successfully. However, these sex-reversed females will usually lack sperm ducts so they do not have the ability to extrude milt from the gonopore (Feist et al., 1995). In order to get hold of the milt the fish has to be killed or a hypodermic needle can be used to sample the milt.

During the first year of growth gynogenetic offspring have high mortality and reduced growth rate compared with normal diploids (Refstie et al., 1982). The mortality is also increased in later ages but ceases as the fish reaches sexual maturation. However there is a great variation in mortality and growth between gynogenetic "families" produced from single females.

A certain number of gynogenetic females may have defective ovaries and some may have traits of intersexuality. However, in many gynogenetic females the reproductive traits are normal. With the high mortality and low frequency of normal broodstock, it is not easy to build up inbred lines for continuous production of large number of hybrids for the industry by crossing inbred lines. For commercial production to be viable, the hybrids need to be of proven superiority in terms of their production traits compared with other available genetic material.

In some experiments it is of importance to have a supply of animals with no or low genetic variation. If gynogenese is repeated for several generation the animals will become completely homozygotes with a coefficient of inbreeding approaching F=1.00. The production and use of inbred lines for research has been discussed by Komen (1990) and Bongers (1997).

18.3 Androgenesis

While the sperm chromosomes are made inactive in gynogenesis, the egg chromosomes are inactivated in androgenesis. A difficulty with androgenesis is that ionising radiation may damage the content of the egg and thus reduce its viability. Thorgaard et al. (1990) produced androgenetic rainbow trout by suppression of first cleavage of eggs with the irradiation and fertilization with sperm from tetraploid males, which gave better results than using sperm from diploid males. However, survival rate was rather low.

One advantage of androgenesis is that homozygous lines with high recovery of genotypes can be produced from cryopreserved sperm. Gene banks of valuable strains can then be created with the cryopreservation of sperm, followed by androgenesis and mating between the androgenetic offspring when they reach maturity.

18.4 Triploidy

Early sexual maturity is a problem in commercial farming as it is often associated with rapid deterioration of flesh quality, higher mortality and reduced growth. A sterile triploid fish will not start producing gonads and therefore should be able to continue to grow undisturbed by the process of sexual maturation and might be slaughtered at any time, Figure 18.3. For some species as tilapia reared in earthen ponds overstocking due to reproduction might reduce the yield and a production based on sterile triploids might have advantages under such circumstances. Sterile farmed fish also have the advantage that any escapees will not mate with wild native stocks and thereby influence the genetic constitution of the wild stock.

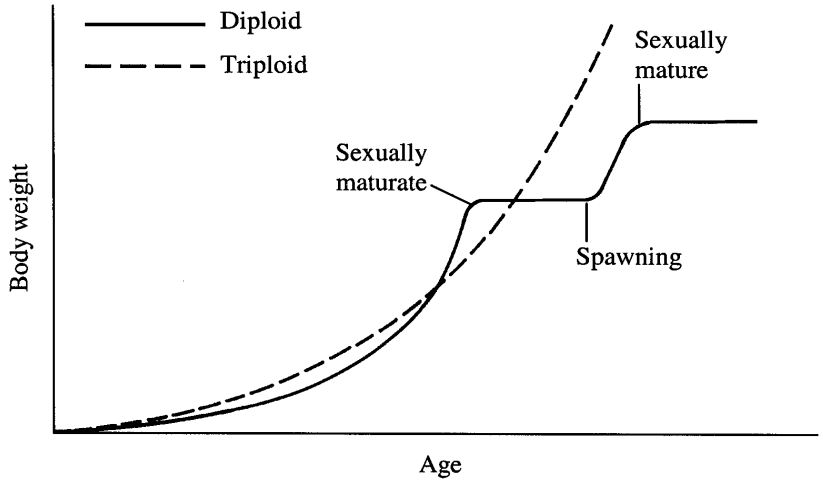


Figure 18.3 Growth curves for diploid and triploid fishes.

18.4.1 PRODUCTION OF TRIPLOIDS

The same technique can be applied to produce triploids as for gynogenesis with the exception that normal and untreated sperm is used. Fertilization is performed by mixing

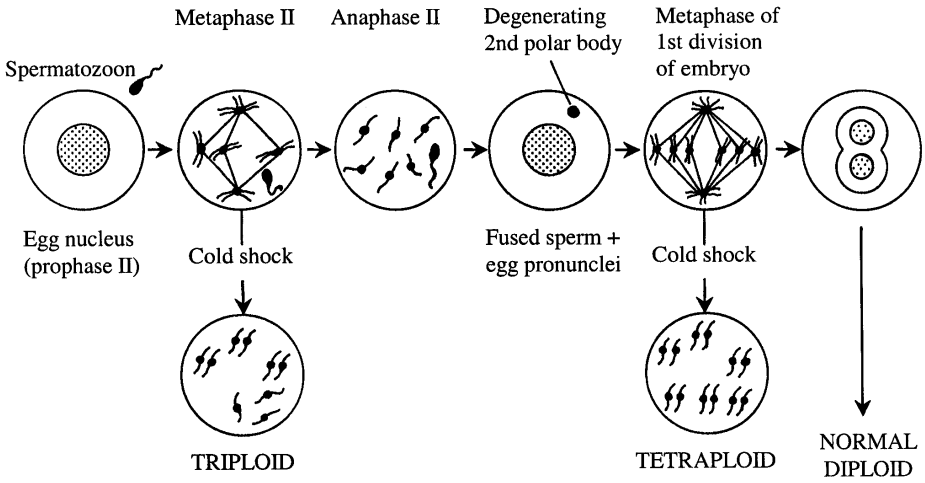


Figure 18.4 Diagrammatic representation if induced polyploidy. Reproduced from Purdom (1993) by permission of Springer Science and Business Media.

egg and sperm. In order to prevent cell cleavage and degeneration of the second polar body, the fertilised eggs are given a temperature shock or exposed to hydrostatic pressure shortly after fertilization. This is illustrated in Figure 18.4. The yield of

triploids is the product of survival of eggs after treatment and the frequency of triploids among survivors, Table 18.1. By varying the temperature and duration of the shock, yields from 0 to 75.5 % were obtained in Atlantic salmon (Johnstone, 1985). The best yield was obtained by applying heat shock of 26 ^{o}C for 20 minutes starting the shock immediately after fertilization. The optimum condition for rainbow trout were similar to those for salmon, but the maximum triploid yield were lower compared with those for salmon (Purdom, 1993).

Table 18.1. Survival to hatching and ploidy status for Atlantic salmon ova following various heat shock treatment. Reproduced from Johnstone (1985) by permission of Elsevier.

Experi Ment No.	Temp. of shock, ^{0}C	Duration of shock, min	Initiation of shock, min after fertilization	Survival to hatch, %	Triploid Rate, %[a]	Triploid yield, %
1	30	4	20	72	30	21.6
	30	6	20	73	50	36.5
	30	8	20	72	82	59.0
	30	10	20	63	92	58.0
	30	12	20	67	100	67.0
	30	14	20	56	100	56.0
2	30	10	0	0	-	-
	30	10	10	50	100	50.0
	30	10	20	67	100	67.0
	30	10	30	73	84	61.3
	30	20	10	0	-	-
	28	20	10	42	92	38.6
3	26	10	0	67	54	36.2
	26	20	0	75	94	70.5
	26	10	10	83	2	1.7
	28	10	0	68	96	65.3
	28	10	10	76	90	68.4
4	32	4	20	69	96	66.2
	32	6	20	0	-	-
	32	8	20	0	-	-
	32	10	20	0	-	-
5	28	10	45	70	16	11.2
6	30	5	35	65	6	3.9

[a] Based on a sample of 50 in each case

Holmefjord (1986) studied the optimal duration of heat shock to produce gynogenesis and triploidy in Atlantic salmon and rainbow trout. He found that when the heat shock was started 10 minutes after fertilization, maximal yields of triploids were obtained at 24 to 26 ^{o}C for both species. Frequencies of triploids were between 90 and 100%. Optimal duration of shock for Atlantic salmon was 24 to 36 minutes at 26 ^{o}C and 32 to 48 minutes at 24 ^{o}C. The corresponding values for rainbow trout were 16 to 24 minutes

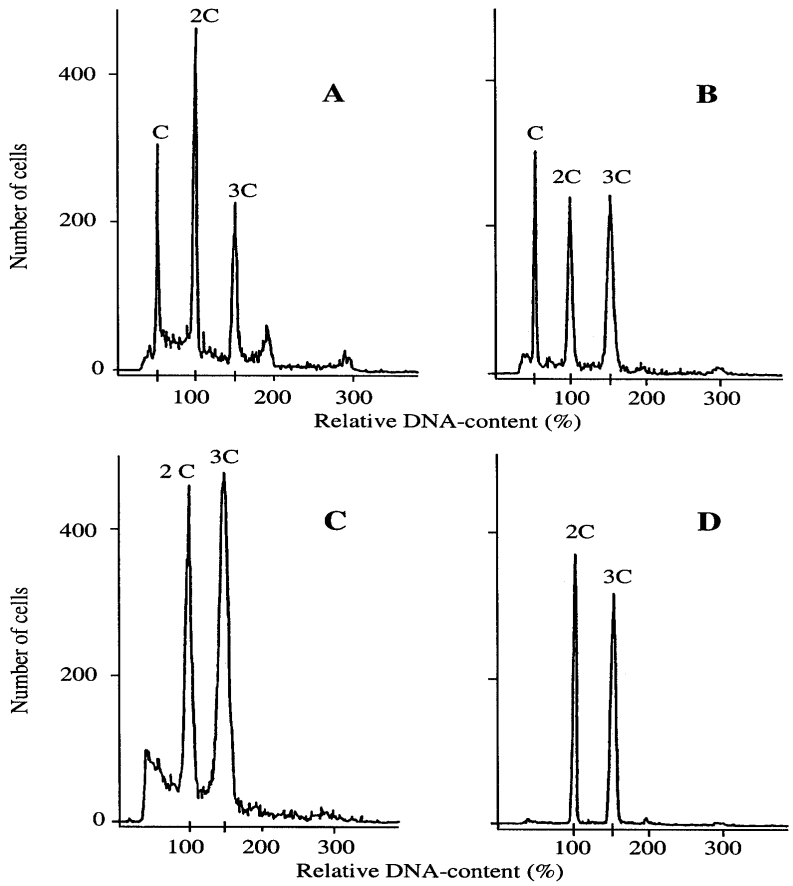
at 26 °C and 24 to 64 minutes at 24 °C. This result shows that many combinations of temperature, duration of shock and time from fertilization to start of shock will produce triploidy. The result also shows that 100% triploids are difficult to produce and some diploids are expected to occur together with the triploids. It is quite common with increased mortality of eggs after heat shock and there is indication that treatment with hydrostatic pressure gives higher survival of eggs when compared with heat shock. Hydrostatic pressure is the most common method for commercial production of triploid salmonids (Lou and Purdom, 1984a). In tilapia Hussain et al. (1991) produced triploids by hydrostatic pressure, heat and cold shock. It was concluded that since production of large number of triploids would require extensive facilities and manpower, production of triploid tilapia would be uneconomic in most situations.

Large scale production of triploid oysters (*Crassostrea virginica)* is reported by Allen and Bushek (1992). Triploidy was induced by inhibiting the second polar body with cytochalasin B. Mean survival of 48 hours old larvae was 22%. The percentage of triploids ranged from 58 to 93 and mean triploidy was 79%. They concluded that triploid *C. Verginica* can successfully produced in high percentages.

18.4.2 METHODS TO DETERMINE PLOIDY

Four different methods have mainly been used to determine triploidy:

1. Chromosome count is based on samples of cells taken from embryos or hatched fry or larvae. The cells are cultivated and stained for later counting of chromosome number. This method is very laborious and it is not possible to count as many cells and individuals which is necessary to get reliable frequencies of triploids
2. Nuclear volumes can be used to identify triploids because there is a high correlation between chromosome number and nuclear volume. It is recommended to use red blood cells for this purpose (Refstie, 1981)
3. Measure of DNA content of individual cells has been done by microfluorimetry or microdensitometry (Gervai et al., 1980)
4. Flow cytometry involves staining cell suspensions with a fluorescent dye, binding specifically to DNA. The amounts of fluorescence from single cells are measured with flow cytometer, which give the frequency distribution of DNA-content per cell. The ploidy can be measured at all developmental stages of cells. Blood samples give ordinary the best results. Figure 18.5 show DNA-histograms for haploid, diploid and triploid Atlantic salmon. This is a fast and accurate method (Thorgaard et al., 1982; Allen,1983; Holmefjord, 1986) while the three methods described first are quite inaccurate and more laborious to determine ploidy (Thorgaard et al., 1982)
5. A count of alleles for microsatellite or other loci could also be used



Stage	Day x degrees C	CV	C : 2C : 3C
A Early embryo	(50-200)	2.1%	51 : 100 : 152
B Eyed egg	(250-600)	2.5%	54 : 100 : 153
C Hatched larvae	(500-600)	4.0%	100 : 148
D Blood cells	(600> d^0)	1.9%	100 : 151

Figure 18.5 DNA-histograms of Atlantic salmon at four different developmental stages (days x degrees C). Peaks: C = haploid, 2C = diploids, 3C = triploid. Reproduced from Holmefjord (1986) by permission of Ivar Holmefjord.

18.4.3 SEXUAL MATURATION OF TRIPLOIDS

Production of triploids is of interest mainly because the fish become sterile. Sterility with triploidy is a result of the disruption caused by three sets of chromosomes in meiosis. In triploid males of plaice Lincoln (1981a) showed that process of spermatogenesis was not seriously affected. Jonasson (1984) found that spermatogenesis appeared normal in triploid rainbow trout and that all stages of cells were produced except spermatozoa. The only difference compared with diploids were that spermatogenesis had not proceeded as far as producing spermatids in some of the

triploid fishes. The weight of gonads 3-4 months prior to sexual maturation was similar in diploid and triploid males. Jonasson (1984) also found that appearance of the testis in triploid males were abnormal. They appeared as a blue-grey and jelly and were often irregular in form compared to diploid testis which were grey coloured and regular in form. Triploid males did not produce motile spermatozoa and no fertilization of eggs was obtained. The triploid males developed secondary sexual characters and the flesh quality was deteriorated in the same way as for maturing diploid males.

Lincoln (1981b) showed that the ovarian development in plaice was almost entirely suppressed in triploid females. In rainbow trout Jonasson (1984) showed that oogenesis in triploids females seemed to be abnormal. The structure of the ovary was normal. Germ cells were recognised both as primary and secondary oogenias. However, oocytes were rarely found. One triploid female had a very undeveloped gonad with primary oogonias present. In this study gonad weight of triploid rainbow trout females was less than 1 gram.

According to Hussain et al. (1995) triploid and diploid tilapia males showed no significant difference in, endocrine profiles or secondary sexual characteristics but the triploids were gametically sterile. Triploid females were found to be functionally and endocrinologically sterile.

18.4.4 TRIPLOIDS FOR FISH FARMING

Male triploids, which develop all the sexual characteristics of diploid males besides being sterile, have no advantage in fish farming. However, females might have advantage since they do not develop gonads and continue to grow while diploids become sexual mature, loses quality and stop growing. Therefore it might be of interest for commercial production to produce all-female triploids (Quillet et al., 1991). This can be done in species where females are the homogametic sex, by producing phenotypic males from genetic females and then use sperm from these sex-reversed females for fertilization of eggs before triploidization. Genetic females can be sex-reversed by adding male hormone to the feed (17 α-methyl-testosterone) during first feeding period. The disadvantage with all-females will be that in some species females grow slower than males, which is the case in salmonids and tilapia while it is the opposite in Atlantic halibut and turbot.

Body weight of triploid rainbow trout 3-4 months prior to sexual maturation averaging 3 kg was found to be similar to body weights in diploids (Jonasson, 1984). Other reports have found that diploids increase the growth rate prior to maturation compared with triploids as illustrated in Figure 18.3. This figure shows the advantage of triploids which continue to grow while the diploids stop growing and continues with sexual maturation. In salmonids triploids only have advantages if sexual maturity occurs before the fish has reached market size.

Galbreath and Thorgaard (1995) found that all- female triploid Atlantic salmon grow significantly slower after one year in sea water compared to diploids. They concluded

that it appear that all-female triploid Atlantic salmon require a lager size to complete smoltification.

Flesh quality in triploid and diploid rainbow trout reared in seawater cages was studied by Bencze (1988). The fish used had grown faster than diploids and most of them were between 4-5 kg. There was no significant difference in chemical content (Fat % and protein %), shape and flesh colour. Sensory judging after smoking and boiling found only small differences between triploid and diploid fish and the differences were too small to be of economic importance for industry concerning quality

According to Hussain et al. (1995) triploid tilapia were not significant different from diploid females in growth rate or proximate composition while in Sydney rock oyster Nell et al. (1994) reports that triploid on average were 41 % heavier than their diploid siblings after 2.5 years of growth. The triploid oysters also maintained higher dry meat weight and higher condition index values when compared with their siblings. Similar results were reported for Pacific oysters (Guo et al., 1996). They crossed diploids and tetraploids with survival comparable to level for diploids and 13-51% larger than their diploid counterparts.

18.5 Production of YY stocks

In some farmed species males grow faster than females and an all- male population might be an advantage in commercial production. When the male is the heterogametic sex all- male population can be achieved by production of YY males and backcross to normal XX females. When the YY males are viable this can be attained by two practices.

Androgenesis will produce 50% females (XX) and 50% YY males, and is an effective way to produce YY males when the male is the hetrogametic sex. However, androgens are rather difficult to produce. Androgenetic diploidy can only be induced by suppression of first cell division, which is a difficult time to manipulate the embryo.

Mair et al. (1997) describe an alternative method for Nile tilapia. They feminized genetic males to phenotypic females by administration of β-estradiol in early life stage. Eggs from the feminised genetic males were fertilized with sperm from normal males (XY). This gave 25% females (XX), 50% normal males (XY) and 25% YY males. Viable males were progeny tested and YY males were identified. When YY males were identified an YY line could be maintained by feminising some of the YY males. A mass production of YY males could then be accomplished through YY x YY mating, eliminating the need for time-consuming progeny testing to discriminate between XY and YY male genotypes.

The overall sex ratio from YY tilapia males mated with XX females ranged from 95% male to the majority being male. YY male genotypes proved to be as viable and fertile as normal XY males (Scott et al., 1989; Varadarjaj and Pandiana, 1989; Mair et al., 1991; Hussain et al., 1994). The progeny of the YY Nile tilapia males have been

comprehensively evaluated in on-station and on-farm trials in the Philippines. Results from on-station trials indicate that genetic modified tilapia (GMT) have benefits under culture, increasing yields by up to 58% compared to mixed sex tilapia of the same strain (Mair et al., 1995).

18.6 Tetraploid

The main interest in producing tetraploid fish is to use them as parents to cross with diploids in order to produce triploids. A procedure for production of tetraploid fish is illustrated in Figure 18.4, basically following the same procedures as for producing triploids. The cold shock is, however, applied at a later stage than for production of triploids, namely at metaphase of 1st. division of embryo. Chourrout (1984) and Myers et al. (1986) were some of the first researchers who succeeded to produce tetraploid rainbow trout.

Tetraploid fish are produced by fertilization of eggs by normal sperm and then suppression of first egg mitosis (Thorgaard et al., 1981; Chourrout, 1982b; Chourrout 1984; Myers et al., 1986; Fosil and Chourrout, 1992; Diter et al., 1993) or trough interspecific hybridization (Liu et al., 2001). In mussels blocking first mitosis in fertilized eggs or fertilization of eggs from triploid females by normal haploid sperm and then inhibiting the first polar body have produced tetraploids (Yang et al., 1997; Yang et al., 1999; He et al., 2000; Supan et al., 2000).

18.7 CONCLUSION

External fertilization in fish and shellfish opens interesting methods for genetic engineering. During the last 30 years researchers in many countries have showed interests in this topic and explored new methods to be used in aquaculture production. The most promising aspects are:

- Production of sterile fish is of greatest interest for species becoming sexual mature before market size like salmonids and tilapia
- Production of mono sex animals is of particular interest in species where one sex grow faster than the other. In some species male grow faster than females like tilapia and salmonids while the females grow fastest in Atlantic halibut *(Hippoglossus hippoglossus),* hake *(Merluccius merluccius)* and in particular in angler fish *(Lophius piscatorius)*
- Production of highly inbred fish in one or a few generations is interesting for utilizing non-additive genetic variation in crossbreeding programs. The methods most used in aquaculture are production of triploid rainbow trout and all-male tilapia. Anon (2004) report that in Great Britain they laid down for hatching 89 % all female diploid eggs, 8 % triploids and only 3 % mixed sex diploids in the year 2002
- Robinson (2002) pointed out that there are several questions which should be answered before starting programs for genetic manipulation (ie. Cost/benefit analysis): Is one gender more highly prized by the market, is one gender of higher

production value to producer, is gender determination known for the species or is it mainly environmentally induced, is clean green market perception/sales pitch going to be jeopardized by manipulation etc.?

19. MODERN BIOTECHNOLOGY AND AQUACULTURE

BEN HAYES AND ØIVIND ANDERSEN

19.1 Introduction

“Biotechnology can be broadly defined as "using living organisms or their products for commercial purposes." As such, biotechnology has been practiced by human society since the beginning of recorded history in such activities as baking bread, brewing alcoholic beverages, or breeding food crops or domestic animals” (Betsch 1994).

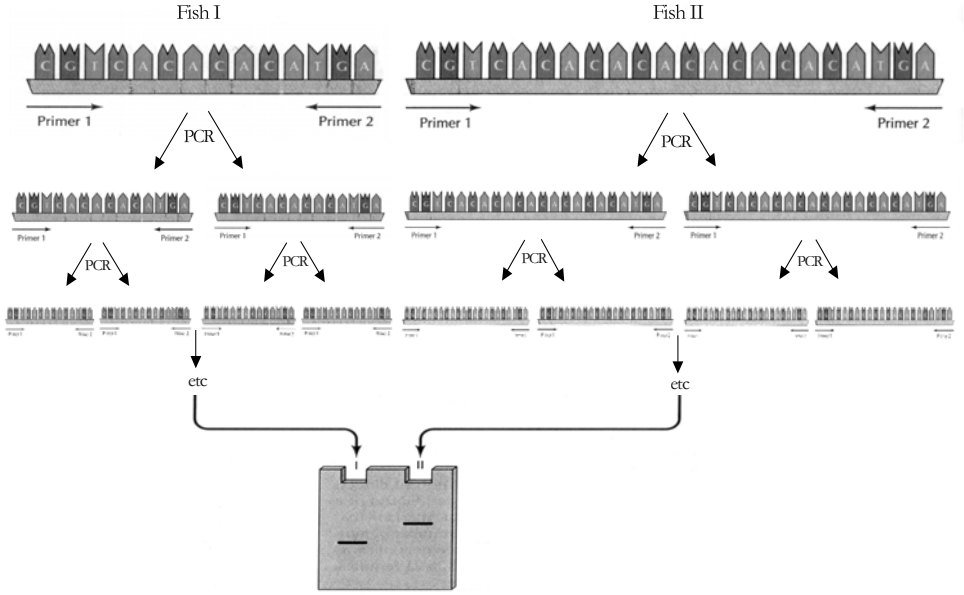


Figure 19.1 Detection of microsatellite markers by PCR (polymerase chain reaction). This microsatellite consists of tandem repeats of the dinucleotide CA. The number of repeats in such microsatellites varies with the individual (e.g. four repeats in fish I and eight repeats in fish II). These variations can be detected by PCR amplifying the repeated DNA fragment with the flanking primers 1 and 2. This process yields amplified DNA products of different lengths as visualized by gel electrophoresis.

Modern biotechnology is concerned with the use and manipulation of the DNA molecules, the genetic code, of living organisms. The emergence of modern biotechnology began with the discovery of a heat stable DNA polymerase enzyme, which produces many copies of small amounts of DNA in the Polymerase Chain Reaction (PCR) (Figure 19.1). With this

T. Gjedrem (ed.), Selection and Breeding Programs in Aquaculture, 301–317.
© 2005 *Springer. Printed in the Netherlands.*

technique selected proportions of the genome, including genes, can be rapidly amplified to amounts suitable for further analysis. Once the DNA sequence is available, the challenge is to use this new information in the genetic improvement of aquaculture species.

The two areas of modern biotechnology, which will probably have the most significant impact on genetic improvement of aquaculture species are DNA markers, and transgenics. In this chapter we investigate potential applications of DNA markers in genetic improvement programs for aquaculture species, including the type of markers available, construction of linkage maps, and accelerated genetic improvement with marker assisted selection (MAS). We also discuss the potential for the application of transgenics in aquaculture, and other emerging technologies.

19.2 DNA markers

A DNA marker is an identifiable physical location on a chromosome whose inheritance can be monitored. Markers can be expressed regions of DNA (genes) or more often some segment of DNA with no known coding function but whose pattern of inheritance can be determined (Hyperdictionary, 2003). If there is no knowledge of these markers for a target species, it is usually possible to discover them rapidly and cheaply using the PCR techniques. For aquaculture species, three types of markers are currently in use. The value of each type of marker depends on the application for which they are intended, their ease and cost of discovery and genotyping, and their level of information content (how many different alleles are present in the population at the marker locus and the frequency of these alleles).

Amplified fragment length polymorphisms (AFLPs) are the result of PCR amplification of a subset of genomic fragments after the DNA has been cut with restriction enzymes and PCR primer binding sites have been ligated to the ends. These amplified fragments are then to be analysed using gel electrophoresis, which generates bands to be compared as polymorphisms. These polymorphisms are typically inherited in a Mendelian genetics fashion, enabling their use for typing, identification and mapping of genetic characteristics. AFLPs are dominant markers, so it is not possible (without advanced software) to distinguish between the heterozygote and the dominant homozygote. Consequently, their information content is rather low. However, AFLPs can be generated and scored very cheaply.

A microsatellite marker is usually a di-nucleotide repeat, for example an AT or CG repeated many times. The alleles are scored according to the size, in base pairs, of the amplified DNA fragments containing the repeated DNA motif. For example, a heterozygous animal may have the genotype 280 282. They have a number of desirable properties, including high polymorphism (many alleles in the population), and consequently high information content, and ease of amplification. Detection of microsatellites is more difficult than detection of AFLP markers (see Figure 19.1), and they are more expensive to genotype. Hundreds of

polymorphic microsatellites have already been identified in many fish species of economical value, but the numbers of potential microsatellite loci present in the fish genome are likely to be large. Specifically, the number of AC dinucleotide repetitions in the Atlantic cod genome was estimated to be in the order of 2.34×10^5, or one every 7,000 base pairs (Brooker et al., 1994). Microsatellites are generally found in the non-coding part of the genome, and so have no effect on gene expression.

One difficulty with micro-satellite markers, which can be a problem in gene mapping, is that they may not exist in sufficient density in the genome to be very closely linked to genes affecting important traits in aquaculture production. An alternative type of marker is the Single Nucleotide Polymorphism, or SNP. These markers are a single base pair substitution at a known site in the genome:

SNPs have the advantage that they occur very frequently throughout the genome, approximately 1 every 1000 bases (0.005cM). SNPs can be in either non-coding DNA or in a specific gene. In the latter case, a SNP in the coding sequence may be a missense mutation causing phenotypic variation. The disadvantage of SNPs is that they are not as informative as microsatellites since they have a maximum of two alleles. Generally about five SNPs are required to give the same amount of information as a single microsatellite (eg. Glaubitz et al., 2003). SNP detection involves comparing sequences from multiple animals for base substitutions. The cost of genotyping SNPs can be much lower than for microsatellites; high throughput technology such as mass arrays allows a low cost for genotyping large numbers of animals for large numbers of markers (Vignal et al., 2002).

19.3 Linkage maps

Once DNA marker loci have been discovered and are readily genotyped, they can be placed on a linkage map. This requires allocating the loci to chromosomes and ordering them along the chromosome. By genotyping a large number of offspring from a number of parents, it is possible to determine which marker loci segregate together, and therefore must be on the same chromosome. It is also possible to determine the distance between the loci on the chromosome. The simplest measure of distance is the fraction of recombinations, or cross overs between the maternally inherited and paternally inherited chromsomes during gametogensis, between two loci per meiosis. The difficulty with this measure is that only odd numbers of recombinants between loci can be observed, as double (and other even numbers) of recombinants between two loci will result in the same marker allele configurations as no recombinants. A consequence of this is that the observed

recombination fraction as a measure of distance is not additive, so the recombination rates between loci A and B and B and C cannot be added to obtain the distance between loci A and C, creating difficulties in the construction of linkage maps with large numbers of markers. An alternative is to use a mapping function. Mapping functions attempt to predict the number of crossovers, also called the map distance (**m**), from the observed recombination fraction (c), accounting for the fact that double and other even numbers of recombinants cannot be observed. The units of **m** are Morgans (M), or centi-Morgans (cM). For example, on average a chromosome of length 1 M would be experience one crossover per meiosis. One of the simplest mapping functions was derived by Haldane (1919), which assumes crossovers occur randomly and independently over the entire chromosome:

$$\mathrm{m} = -\frac{\ln(1-2\mathrm{c})}{2} \tag{19.1}$$

This mapping function does not account for interference, where the presence of a crossover in one region affects the frequency of crossovers in other regions. The Kosambi (1944) mapping function does take some account of interference:

$$\mathrm{m} = \frac{1}{4}\ln\left(\frac{1+2\mathrm{c}}{1-2\mathrm{c}}\right) \tag{19.2}$$

A small example illustrates the distance between two loci predicted by the Haldane and Kosambi mapping functions. For two marker loci on the same chromosome, A and B, a female broodstock carries the alleles A1_B1 on the chromosome she has inherited from her mother and A2_B2 on the chromosome she has inherited from her father. She has ten progeny, four receive alleles A1 and B1, four receive A2 and B2, one receives A1 and B2 and one receives A2 and B1. So there are two recombinant progeny, and the recombination frequency is 2/8=0.25. Using the Haldane mapping function, the distance between the two loci is $-\frac{\ln(1-2*0.25)}{2} \approx 0.35$ Morgans. If we use the Kosambi mapping function, the distance between the two loci is $\frac{1}{4}\ln\left(\frac{1+2*0.25}{1-2*0.25}\right) \approx 0.27$ Morgans. In general, at small recombination frequencies, the observed recombination fraction, the Haldane mapping function and the Kosambi mapping function predict very similar map distances, while with large recombination fractions, the difference between the predicted distances increases, as double recombinants become increasingly likely, Figure 19.2.

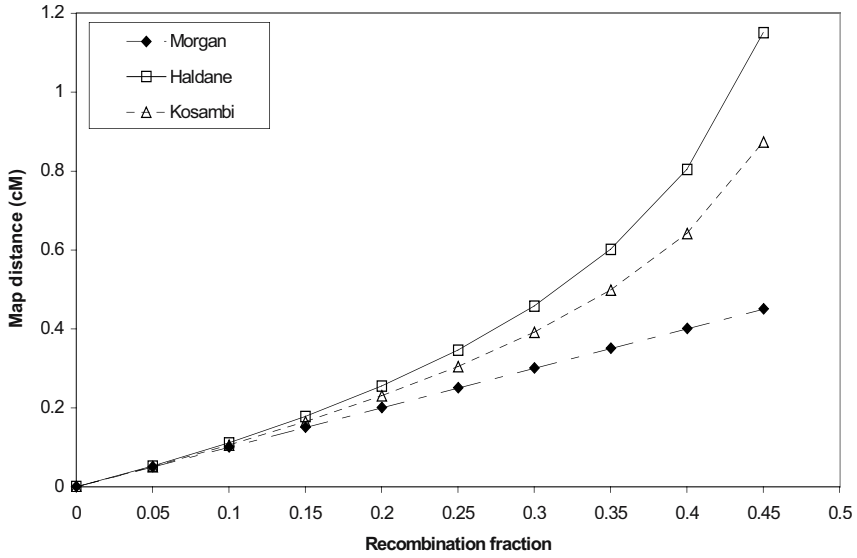


Figure 19.2 Relation between recombination fraction and map distance for Morgan's, Haldane's and Kosambi's mapping function.

Both the Kosambi and Haldane mapping functions are widely used in the construction of linkage maps. Figure 19.3 is an example of a linkage map in *Penaeus monodon* constructed with the Kosambi mapping function.

Many of the aquaculture species show dramatic differences in recombination rates between the sexes. For example, in zebrafish, the recombination rate in females has been estimated to be 2.74 times higher than in males, (Singer et al., 2002). In salmonids, the difference is even more dramatic, with recombination in males approximately eight times greater than in females of Atlantic salmon (Moen et al., 2004b) and 3.25 about 6 times greater in the genus of *Oncorhynchus* (Sakamoto et al., 1999). This leads to dramatic differences in the length of the linkage map, for example Figure 19.3. It is therefore important to consider possible sex differences when constructing linkage maps for aquaculture species.

The amount of DNA marker information available for the various aquaculture species generally reflects their economic or scientific importance. For example, the linkage map for zebrafish, used as a model organism in disease and developmental studies, has many thousands of markers (Woods et al., 2000)) At the time of writing, linkage maps, often with a combination of AFLP markers and microsatellites, were available (at least) for tilapia (Kocher et al., 1997; Kocher et al., 1997; McConnell et al., 2000), rainbow trout (Young et al., 1998; Sakamoto et al., 1999; Nichols et al., 2003), Atlantic salmon (Moen et al., 2004b),

common carp (Sun and Liang, 2003), Kumura prawn (Moore et al., 1999) and *Penaeus monodon* (Wilson et al., 2002). The AFLP markers are used to complement the microsatellite markers as frequently the number of microsatellite markers available for a particular species is low. Linkage maps for many other species will be published in the next few years.

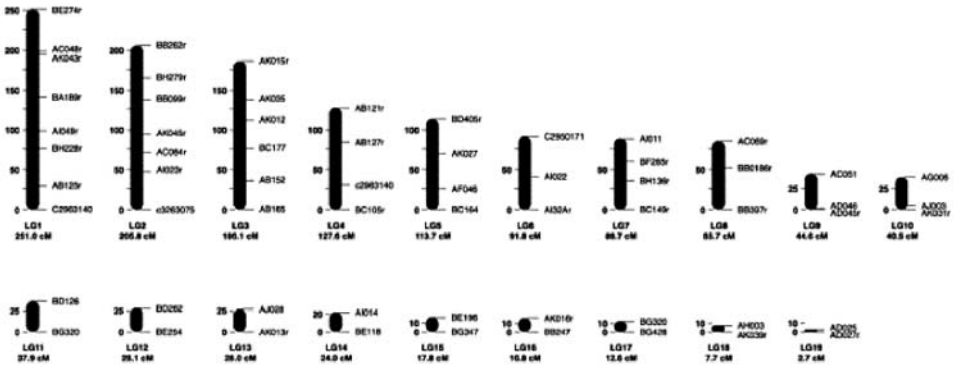


Figure 19.3. Linkage map for *Penaeus monodon.* The length of each chromosome represents the genetic length (cM) of that chromosome. The female (top) and male (bottom) maps are presented separately. Reproduced from Wilson et al. (2002) by permission of Elsevier.

19.4 Linkage mapping and QTL analysis

19.4.1 DETECTION OF QUANTITATIVE TRAITS LOCI

Once a linkage map has been created, it can be used to determine if variation at the molecular level (allelic variation at marker loci along the linkage map) is linked to variation in phenotypes. If this is the case, then the marker is linked to, or on the same chromosome as, a quantitative trait locus or QTL which has allelic variants causing variation in the quantitative trait. To illustrate the principle of QTL mapping, consider an example where a

particular sire has a large number of progeny. The parent and the progeny are genotyped for a particular marker. At this marker, the sire carries the alleles 172 and 184, Figure 19.4. The progeny can then be sorted into two groups, those that receive allele 172 and those that receive marker 184 from the parent. If there is a significant difference between the two groups of progeny, then this is evidence that there is a QTL linked to that marker.

For a given size of effect of QTL, the contrast between the two groups will be larger the closer the marker is to the QTL on the chromosome, as recombination reduces the strength of the QTL marker association. The size of the QTL effect, the average effect of a QTL allele substitution, is estimated as $\alpha = d/(1 - 2r)$, where d is the contrast between the two groups of progeny receiving each allele, and r is the distance between the QTL and the marker. With only one marker, it is not possible to estimate both r and α, as there is only one known parameter in the equation (d). The solution is to use multiple markers in a process known as interval mapping (Lander and Botstein, 1989), where information from all the markers on a chromosome is used simultaneously to estimate the position of the QTL and the size of its effect.

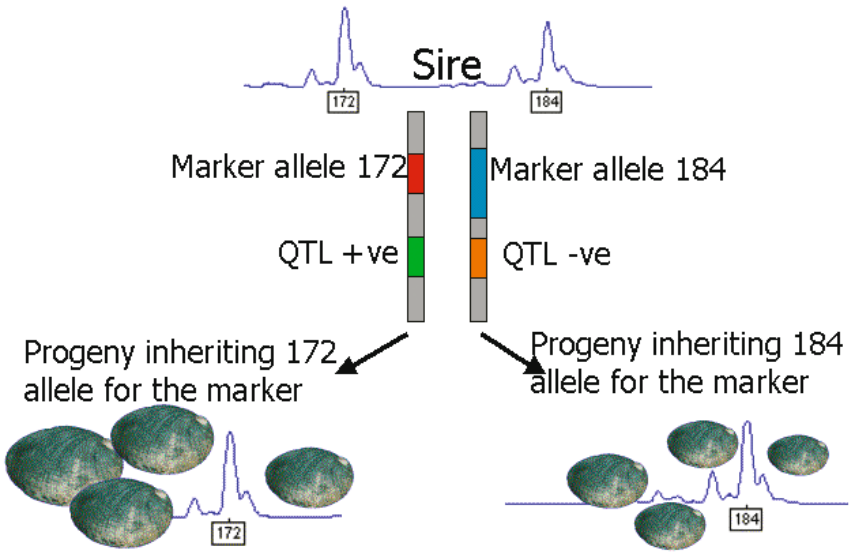


Figure 19.4. Principle of quantitative trait loci (QTL) detection, illustrated using an abalone example. A sire is heterozygous for a marker locus, and carries the alleles 172 and 184 at this locus. The sire has a large number of progeny. The progeny are separated into two groups, those that receive allele 172 and those that receive allele 184. The significant difference in the trait of average size between the two groups of progeny indicates a QTL linked to the marker. In this case, the QTL allele increasing size is linked to the 172 allele and the QTL allele decreasing size is linked to the 184 allele (Figure courtesy of Nick Robinson, AKVAFORSK).

Using interval mapping, and provided sufficient markers are available, the entire genome can be 'scanned' for QTL. Support for existence of a QTL is evaluated at regular intervals along the genome. Information from the markers flanking each interval is used. Two statistical methodologies are widely used in genome scans to evaluate the presence or absence of QTL, maximum likelihood (ML) and regression. Maximum likelihood (ML) methods calculate at each putative QTL position, the log of the ratio of the maximum of the likelihood function under the null hypothesis of no segregating QTL to the maximum of the

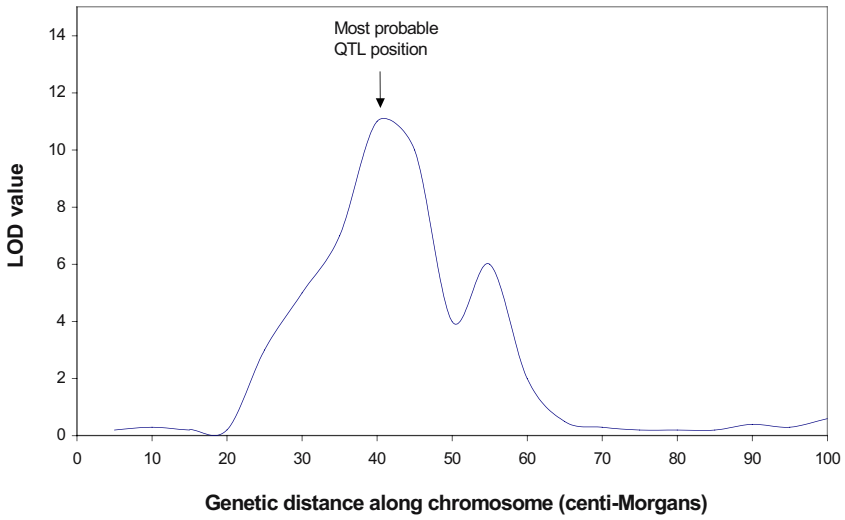


Figure 19.5. A possible likelihood of odds (LOD) profile along a chromosome from a genom scan. The most probable position of a QTL is indicated.

likelihood function under the alternate hypothesis of a QTL segregating at that position, given the marker and phenotype data. The result is a likelihood of odds (LOD) score for each putative QTL position, which creates a LOD profile across the chromosome (Figure 19.5). The point at which the LOD profile is the highest is the most probable position of the QTL. An alternative to ML is to use regression methodology. In this approach, the probability that progeny received particular QTL alleles from their parents are calculated from flanking marker information, and the phenotypes of the progeny are regressed onto these probabilities (Knott and Haley, 1992). Results from the regression method have been shown to be similar to the ML method, and takes much less computer time. For more detailed description of statistical methodologies for QTL mapping, see Lynch and Walsh (1998).

Before embarking on a QTL mapping experiment, it is very important to have some idea of the power of the proposed experiment to detect QTL. Results from QTL mapping experiments in livestock indicate that for most traits, even the QTL with the largest effect will have comparatively small effects (eg. Carlborg et al., 2003), so powerful mapping experiments will be required. The power to detect and locate QTL in linkage experiments depends on two factors, the accuracy with which QTL effects are estimated, determining the ability of the experiment to separate the QTL effect from environmental and other effects, and the total number of recombination events in the experimental animals. The more recombination events there are in the experiment, the more precise the estimation of the location of QTL.

The experimental population, which will be used in a QTL detection experiment, depends on the reproductive biology and existing lines of the species in question. In fish, backcross and F_2 populations can often be created from the intercrosses of different lines. As alternate alleles for both marker and QTL loci may be fixed, or at markedly different frequency in the alternate breed or lines, both types of loci may be segregating in the F_2 or backcross progeny. These populations are a powerful resource for QTL mapping. F_2 populations have been used as a resource for QTL mapping in number of fish species, for example Tilapia (Cnaani et al., 2003) and rainbow trout (Ozaki et al., 2000). An alternative design is to utilise the large full-sib and half sib families, which exist in the breeding programs from a number of species. A genome scan of the salmon genome for disease resistance QTL used this type of population (Moen et al., 2004a).

The power of a mapping experiment to detect QTL using F_2 populations from crosses of inbred lines depends on the number of studied offspring and the size of the QTL effect. Lynch and Walsh (1998) give an expression for the number of F_2 progeny required to detect QTL explaining a proportion r^2_{F2} of the F_2 segregation variance with a power of $1-\beta$:

$$n_{F2} = \left(\frac{1-r^2_{F2}}{r^2_{F2}}\right)\left(\frac{z_{(1-[\alpha/2])}}{\sqrt{1-r^2_{F2}}} + z_{(1-\beta)}\right)^2 \left[1 + (k^2/2)\right] \qquad (19.3)$$

where α is the significance level, and z is the ordinate of the normal distribution at the given level of significance. So to detect a QTL explaining 10% of the total F_2 variance with a power of 90% and with α= 0.05, with completely additive gene action (k = 0), 101 F_2 progeny would be required. If the parental lines are not completely inbred, the power is also a function of the difference in allele frequencies at the marker and QTL loci between the parental lines. The more similar the allele frequencies, the lower the power of the design to detect QTL.

The probability that a QTL mapping experiment detecting a QTL power of a mapping experiment in outbred populations, such as the breeding programs for salmon in Norway,

depends on the heterozygosity of segregating QTL and the number of individuals per family. Martinez et al. (2002) investigated the power of various designs of QTL mapping experiments in fish to detect QTL. They investigate full sib designs, hierarchical designs (one male mated to more than one female) and double haploid designs (see chapter 12?? for a description of double haploids). Their results suggest large full sib families are required if the QTL mapping experiment is to stand any chance of detecting QTL of small to moderate effect, Table 19.1. For example, even with 1000 progeny, in one full sib group, the power to detect a gene with effect of 0.2 is only 0.51 (ie. the QTL is only detected in 51% of genome scans of this size) . If double haploids are used, the size of the full sib families can be reduced significantly while retaining the same power of QTL detection.

Table 19.1. Power of QTL detection for different fish population structures, from simulation. Power (corresponding family size in parenthesis) is optimised over family structure for the double haploid (DH) and full sib (FS) designs, for different population size (No) , QTL effect (a) and residual genetic variation (σ_g^2). The environmental variance was 1 in all cases.

σ_g^2		DH			FS		
		0.0	0.5	1.0	0.0	0.5	1.0
	No.	Power			Power		
0.2	200	0.276 (200)	0.180 (200)	0.129 (200)	0.084 (200)	0.065 (200)	0.053 (200)
	400	0.449 (400)	0.355 (400)	0.276 (400)	0.199 (400)	0.151 (400)	0.120 (400)
	600	0.573 (300)	0.449 (600)	0.386 (600)	0.319 (600)	0.248 (600)	0.199 (600)
	800	0.660 (400)	0.531 (400)	0.449 (800)	0.423 (800)	0.341 (800)	0.280 (800)
	1000	0.731 (333)	0.607 (500)	0.507 (500)	0.507 (1000)	0.423 (1000)	0.356 (1000)
0.4	200	0.660 (100)	0.531 (100)	0.449 (200)	0.423 (200)	0.341 (200)	0.280 (200)
	400	0.869 (57)	0.750 (133)	0.660 (200)	0.660 (400)	0.594 (400)	0.531 (400)
	600	0.955 (50)	0.870 (100)	0.787 (150)	0.787 (300)	0.711 (300)	0.660 (600)
	800	0.986 (50)	0.936 (80)	0.869 (133)	0.869 (133)	0.806 (400)	0.749 (400)
	1000	0.996 (50)	0.969 (83)	0.923 (111)	0.923 (111)	0.870 (333)	0.817 (500)

19.4.2 STRATEGIES TO MINIMISE THE NUMBER OF GENOTYPINGS

The most significant cost in genome scanning is usually the cost of genotyping animals at the DNA markers. A number of strategies have been devised to reduce the cost of genotyping.

In selective genotyping, only individuals from the high and low phenotypic extremes are genotyped (Darvasi and Soller, 1992). If a half or full sib design is being used, the selective genotyping is done with each family. Selective genotyping exploits the fact that individuals most deviating from the mean are considered to be most informative for linkage, as their genotypes at the QTL can be inferred from their phenotypes more accurately than can those with average phenotypes.

Selective DNA pooling allows even further reduction of genotyping costs (Darvasi and Soller, 1994; Lipkin et al., 1998). The determination of linkage between a marker and a QTL in DNA pooling is based on the distribution of parental alleles among pooled DNA samples of the extreme high and low phenotypic groups of offspring. The concept is illustrated in Figure 19.6.

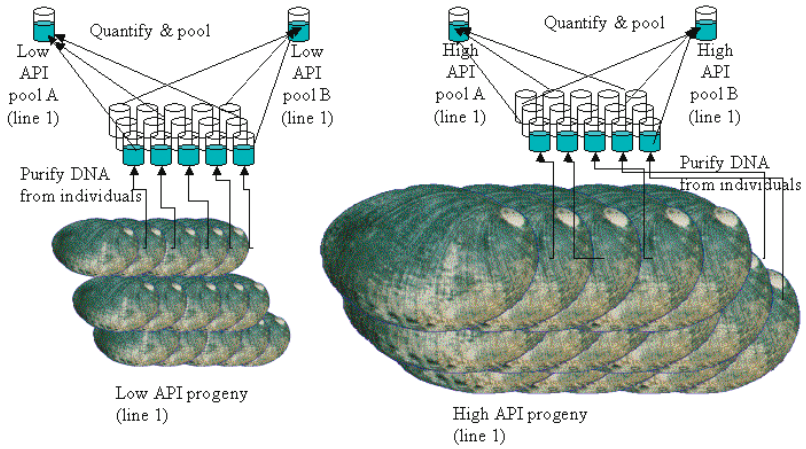

Quantify & pool
Low API pool A (line 1)
Low API pool B (line 1)
Purify DNA from individuals
Low API progeny (line 1)
Quantify & pool
High API pool A (line 1)
High API pool B (line 1)
Purify DNA from individuals
High API progeny (line 1)

Figure 19.6 Design for selective DNA pooling of high and low growth rate abalone progeny within an abalone family line. Two pools of lows and two pools of highs are created for replication. The pools can then be genotyped for DNA markers and the abundance of alleles in each pool measured (Figure courtesy of Nick Robinson, AKVAFORSK).

For a particular sire, if at a microsatellite locus the marker allele 150 is linked to the ncreasing QTL allele (Q), and the marker allele160is linked to the decreasing QTL alleles (q), then we would expect more of the 150 marker in the high pool than and in the low pool, and more of the 160 allele in the low pool than the high pool, Figure 19.7. A DNA pooling experiment generally has high power to detect QTL affecting only the trait from which the pools of high and low phenotype progeny were made. The power to detect QTL for other traits is low, unless there is a substantial genetic correlation between the traits.

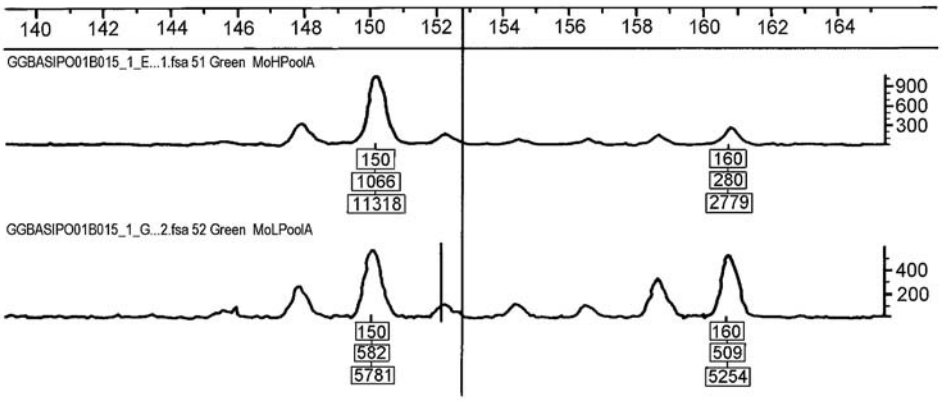


Figure 19.7. Electropherogram of marker allele abundance in high and low phenotype pools of progeny (top and bottom line respectively) from a single sire. The sire has two alleles at the microsatellite marker, 150 and 160 (Figure courtesy of Nick Robinson, AKVAFORSK).

19.5 Results of QTL detection

At the time of writing, the number of detected QTL in aquaculture species was rather small, and confined to tilapia, carp, rainbow trout and Atlantic Salmon. Sun and Liang (2003) reported a QTL for cold tolerance segregating in cross between a common carp and a Boshi carp (*C. pellegrini pellegrini*). Cnaani et al. (2003) reported QTL for cold tolerance and fish size in a F_2 tilapia hybrid. QTL for salinity tolerance (Lee, 2003) and red body colour (Howe and Kocher, 2003) in tilapia have also been reported. In rainbow trout, Ozaki et al. (2000) reported two putative QTL, each with significant effects on resistance to infectious pancreatic necrosis (IPN). Finally, Moen et al. (2004a) reported a QTL with significant effects on resistance to infectious salmon anaemia (ISA) in Atlantic salmon.

19.6 Implementing marker assisted selection in breeding programs for aquaculture species

If DNA markers are close to a genetic mutation or QTL, which has a favourable effect on a quantitative trait, then this information can be used to more accurately select broodstock, which are genetically superior for the quantitative trait, Figure 19.8.

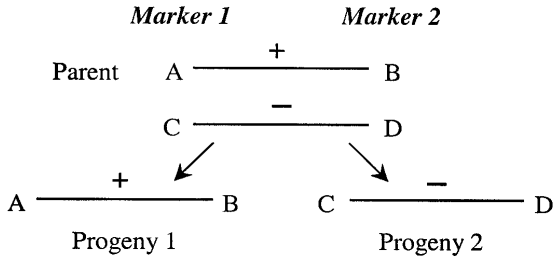


Figure 19.8. Paternal and maternal copies of a small segment of a male broodstock chromosome, and two of his offspring. There is a QTL affecting growth rate between markers 1 and 2. On his paternal chromosome, the male broodstock carries a QTL which increases growth rate (indicated by +), on his maternal chromosome he carries an allele which decreases growth rate.

For marker one, the parent has allele A on his paternal chromosome and allele C on his maternal chromosome. For marker two, the broodstock has allele B on his paternal chromosome and allele D on his maternal chromosome. Between the two markers is a QTL. The broodstock carries a mutation, which increases the quantitative trait on his paternal chromosome. By genotyping his offspring for marker one and marker two, we know they have inherited the paternal or maternal copy of the broodstock's chromosome, and therefore if they have inherited the favourable mutation (ignoring crossovers in this example). The marker information can be used as a tool to select progeny for future breeding. This process is termed marker assisted selection (MAS).

To implement marker assisted selection, phenotypic, pedigree and marker information is combined to estimate marker assisted breeding values, or M-EBVs (Fernando and Grossman, 1988). Broodstock can then be ranked on their M-EBVs and the desired number selected. M-EBVs estimated using a marker bracket, as in Figure 20.8, are more accurate than M-EBVs estimated from only one marker, unless the marker is very closely linked to the QTL (Meuwissen and Goddard, 1996). In practise, using markers which are not closely linked with the QTL is difficult because large families are needed to estimate allele effects accurately and marker-QTL phase information (which marker allele is associated with the favourable QTL allele in each family) must be continuously re-estimated. Such markers are therefore only really practical in the species with a nucleus serving a large population (for example the Norwegian salmon industry). By contrast, a very tightly linked marker or a direct test for the QTL is easy to utilise because the individual breeder can rely on estimates of the QTL allele effects obtained elsewhere.

If accuracy of predicting breeding values is already high without the marker information, the DNA marker data will not add much extra accuracy (Goddard and Hayes, 2002). For instance, for a trait with a heritability $h^2 = 0.5$, a single phenotypic record yields an EBV with accuracy h=0.71. Simulation studies investigating the advantage of MAS with a trait

that is measured on all animals prior to selection (eg growth rate), find only marginal gains from addition of marker data (Lande and Thompson, 1990; Meuwissen and Goddard, 1996). Gains are potentially much larger for traits such as disease resistance or meat quality traits, such as filet colour in salmon. For these traits, the phenotype is not recorded on the individual, rather information comes from relatives of the potential breeding candidate or from a progeny test. In this situation, the breeding candidate can be genotyped for the marker(s) prior to selection, and if the QTL have a large effect, the accuracy with which her or his breeding value is predicted can be considerably increased (Goddard and Hayes, 2002).

One of the advantages of using marker information in aquaculture breeding programs is that it may be possible to identify breeding candidates of high genetic value at an early age. For example, in a salmon breeding program, in the future markers linked to a QTL explaining a large proportion of the genetic variance for filet colour and disease resistance could be available. Then it would be possible to identify and select progeny carrying favorable QTL alleles for these traits at a very early age. This would circumvent the need to rear large numbers of progeny per family for testing.

19.7 Other applications of DNA markers

Selective breeding programs in aquaculture make use of family information, which requires that families are kept separately until the fish are large enough to be physically tagged. This imposes major economic and practical problems, and can induce environmental effects common to full-sibs. Identification of family groups by use of DNA markers has the potential to overcome these problems. Using this technology, fish from different families could be reared together in the same tank even from the egg stage. As the need to keep each family in a separate tank is circumvented, using DNA markers would allow larger number of families to be tested, without increasing the number of tanks, and thus facilitating the use of higher selection intensities without rapid accumulation of inbreeding (Estoup et al., 1998).

Due to their high polymorphism, microsatellite markers are useful for genetic tagging. Microsatellites have successfully been used to empirically reconstruct pedigrees in fish populations with families mixed from hatching (Hallerman and Beckmann, 1988). In an experiment with Atlantic salmon families, using four highly polymorphic microsatellites was sufficient to assign at least 99% of the offspring to the correct pair with 100 crosses involving 100 males and 100 females (Hallerman and Beckmann, 1988). An additional polymorphic microsatellite was required for correctly assigning 99% of the offspring when the 100 crosses were produced with 10 males and 10 females (Hallerman and Beckmann, 1988). This study demonstrated that parental assignment is feasible with the DNA markers currently available in several fish species.

Both the efficiency and costs of microsatellite based pedigree analysis should be considered before this method is included in a breeding program. Practically, the parents and the mixed offspring are genotyped by PCR amplifying the appropriate microsatellite loci from crude DNA extracts from small non-destructively sampled quantities of tissue such as fish scales, mucus, and fin clip. Using this protocol, approximately 2000 fish from a mixture of 500 families can be screened for 10 markers in less than a month allowing 99 % of the fish to be parental assigned (Verspoor pers. com.). Allocation of offspring to families without parental genotypes is also possible, provided enough fish are sampled to obtain sufficient representation of all families, but a larger number of markers are required than for parentage analysis, with probably more than one hundred microsatellites required (Garcia et al., 2002).

The steady decreasing costs of genetic tagging may still not compete with traditional physical tagging. As the genotype information is detached from the individual, genetic tagging implies that the fish has to be retyped each time its performance is evaluated or individuals are selected. Alternately, fish can be physically tagged for re-identification following genotyping and parentage assignment. Thus the cost of implementing DNA markers in selection programs could be considerable. However it should soon be possible to apply recent developments in human genetics such as DNA chips, where many SNP markers can be genotyped simultaneously and cheaply, in the aquaculture species, and this could dramatically reduce the cost of genetic tagging. DNA markers have several other applications within fish management as well, including evaluation of inbreeding levels, stock identification, movement of released or escaped fish and their possible genetic interactions with wild stocks.

19.8 Transgenic fish

The high fecundity of most fish and external fertilisation and embryonic development make them especially suitable for transferring specific genes. Successful gene transfer has been demonstrated into a variety of aquatic organisms by applying different techniques, including microinjection, particle gun bombardment, and electroporation. The transfer of genes into the sperm or directly into the skeletal muscle have become alternatives to the fertilized eggs. The production of transgenic fish is aimed at dramatic improving traits like growth, disease resistance, and environmental tolerance. The nature itself of such quantitative traits, being influenced by multiple genes, makes them difficult to manipulate by gene transfer techniques. However, significant growth enhancement of several fish species has been demonstrated after the introduction of the growth hormone gene as the major gene under the control of a strong promoter element (Melamed et al., 2002). It should be noted that the increased growth rate in the transgenic fish is mainly due to the great amounts of growth hormone produced in large tissues like the liver or gonads, and not the result of injecting millions of gene copies into the fertilized eggs. Surprisingly, the growth of transgenic wild-strain rainbow trout was demonstrated to not surpass that of non-transgenic domesticated strain being selected for fast growth (Devlin et al., 2001). Furthermore, introducing the

growth hormone construct into this fast-growing domestic strain did not cause further growth enhancement. The results indicate that similar alterations of growth rate can be achieved both by selective breeding and by transgenesis, but that the effects are not additive at least not in rainbow trout.

There are several problems to be overcome before transgenic animals can be produced on a large scale. Indeed, in over 90 % of the microinjected eggs, the transgene is not efficiently integrated in the genome at the one-cell stage. The result is highly mosaic transgenic fish and low frequencies of germ line transmission since only the tissues developing from the transformed cell will carry the transgene. Furthermore, the injected DNA integrates at single or multiple random sites in the genome of the recipient embryo, each developing into an unique hemizygous fish. Hence, the establishment of a stable transgenic broodstock will be a costly endeavour requiring several generations.

On the other hand, attempts to combat viral and bacterial pathogens threatening commercial stocks by utilizing DNA vaccines have been promising. This technique is based on the injection of DNA encoding part of the antigen, usually a bacterial outer membrane or viral capsid protein, in the fish muscle. Here the protein will be synthesized, and the production of antibodies against the foreign protein is induced. A significant degree of protection against infectious hematopoetic necrovirus (IHNV) was found in Atlantic salmon after vaccination with a gene construct containing an IHNV glycoprotein (Traxler et al., 1999). Similarly, protection against viral haemorrhagic septicaemia virus (VHS) was induced in vaccinated rainbow trout (Lorentzen et al., 1999). The main disadvantage of these approaches is they require detailed information about the structure, conformation, and the encoding sequence of the pathogen's protein.

An alternative approach to increase the resistance of fish to pathogens is to target the non-specific immune response through use of antimicrobial peptides, which are found in both vertebrates and invertebrates. Short peptides consisting of 30-50 amino acids with strong antimicrobial activity have been isolated from the skin mucus of several fish species, Recently, manipulation of antimicrobial cecropin genes in the Japanese medaka and channel catfish produced strains of transgenic fish resistant to infection by fish bacterial pathogens (Sarmasik et al., 2002; Dunham et al., 2002).

19.9 Future prospects

The applications of molecular techniques in aquaculture are promising, but still somewhat uncertain. While high costs seem to be the only hindrance for widespread application of DNA markers for identification purposes and marker assisted selection, the situation regarding commercial use of genetically modified fish is more complex. Although the potential importance of gene transfer technology is large, a major concern relates to the possible impact, which release or accidental escapes of gene-modified individuals may have

on natural ecosystems. Other controversial aspects are related to animal welfare, food safety and the public perception of gene manipulation in general. To which extent such issues will constrain the future use transgenic animals in applied aquaculture production remains to be seen.

Other technologies are also rapidly emerging which are either being used or are likely to be used in the future in the aquaculture species. For example micro-array technology, where an array of DNA or protein samples are hybridized with probes to study gene expression of a large number of genes simultaneously, can be used to determine which genes are involved in response to disease, expression of meat colour and other important issues. Already, a micro-array containing DNA probes for 3700 DNA sequences or genes is available in salmon, and will be used to compare gene expression in disease challenged and non-challenged fish (Davidson and Koop, 2003). Such technology has the potential to contribute very large amounts of information on the genes and pathways of genes, which affect the economic traits in aquaculture species.

20. GENETIC INTERACTIONS BETWEEN FARMED AND WILD FISH, WITH EXAMPLES FROM THE ATLANTIC SALMON CASE IN NORWAY

HANS BERNHARD BENTSEN AND JØRN THODESEN

20.1 Introduction

Aquaculture is frequently associated with introduction of exotic species or populations and genetic modifications of the farmed organisms. Domesticated, high performing breeds of aquaculture species are mainly developed by a long-term natural selection in captivity and accumulation of genetic gain caused by generations of artificial selection. In addition, several technologies, such as crossbreeding schemes, gynogenesis, chromosome manipulations and gene transfer (among others) are available. These different technologies are usually evaluated based on their potential to make farmed species more efficient, with faster growth and improved feed utilization for example. However, there is a growing public awareness concerning what might happen with the biological diversity of the environment if exotic and/or genetically improved fish escape from aquaculture and interact with wild, local populations. Since no fish farming system can be considered fully secure against the escape of the farmed organisms, available technologies for genetic improvement of fish should also be evaluated according to their potential environmental risks. Because of limited information about the impact of escapees for most aquaculture species, , farmed Atlantic salmon in Norway is used here as a case study to illustrate possible interactions between genetically improved and wild fish.

20.2 Conservation of biological diversity

In 1992, the United Nations (UN) agreed upon a convention to protect the global biological diversity (the Biodiversity Convention). However, concepts such as "diversity of species" and "genetic diversity" are often used as synonyms to "biological diversity" when discussing necessary actions to follow up this convention. Although these concepts are all related to each other, it might be necessary to give a brief definition of each concept before discussing possible environmental impacts of genetically improved fish.

20.2.1 DIVERSITY OF SPECIES

A species is often defined as all individuals that may produce fertile and competitive offspring when mated to each other. This means that genes usually cannot be transferred from one species to another, and that entire genomes will be lost forever if a species is lost. Therefore, conservation of species is the most basic and important effort to secure future biological diversity.

T. Gjedrem (ed.), Selection and Breeding Programs in Aquaculture, 319–334.
© 2005 *Springer. Printed in the Netherlands.*

20.2.2 GENETIC DIVERSITY

Genetic diversity also includes the diversity of populations and individuals within species, and has consequently a wider definition than the diversity of species. Most wild species are divided into natural populations, as well as farmed species in domesticated stocks or breeds, which are more or less genetically isolated from each other. Populations of the same species might become genetically different from each other by chance (i.e. mutations and genetic drift) or due to genetic adaptation to different environments (i.e. natural or artificial selection). If a species is separated into a number of diversely adapted populations, the loss of one population may not always be compensated by the expansion of another population into the free habitat. Loss of populations may therefore threaten the viability and eventually the survival of the species.

20.2.3 BIOLOGICAL DIVERSITY

Biological diversity is caused by both genetic diversity and differences in environmental conditions. Different environmental conditions, such as different climate and supply of nutrients, might create biological diversity between populations even when there is no or very limited genetic difference between them. It follows that biological diversity has a much broader definition than both the diversity of species and genetic diversity, and that conservation of biological diversity might require that environmental diversity be also protected.

20.2.4 LONG-TERM CONSERVATION

The number of species on earth has not been constant or continually increasing over historical time. Studies of fossils suggest that there have been periods with both increasing and decreasing number of species. These periods have sometimes been interrupted by incidents of dramatic reductions in the diversity of species, most likely due to sudden changes in the climate and environmental conditions. Therefore, the long-term goal for conservation of biological diversity cannot be to preserve all species forever, but rather to prevent a severe acceleration of natural changes in the diversity of species because of human activities. In securing long-term conservation of biological diversity, the main focus should be to protect species and, sometimes, populations within species. However, conservation of threatened populations that have a similar genetic constitution as other more viable populations will not contribute much to conserve the genetic diversity and viability of a species. On the other hand, large, viable populations with a long lasting, genetically isolated evolution history may have developed unique genetic adaptations to specific environments and possibly accumulated fitness mutations. Such populations are expected to have a greater importance for maintaining the diversity of specific genetic adaptations and, thereby, the viability of the species in the future. In accordance with this, the concept of stock management of fish species was developed the early 1980-ies (Billingsley, 1981; Ryman, 1981). However, good criteria to identify such critical populations have not yet been developed.

20.3 Introduction of exotic species of fish

The increasing demand for aquaculture products has created a considerable interest in many countries to introduce farming of new profitable species of fish. The result has been the introduction of several farmed species, such as Atlantic salmon and Nile tilapia, in parts of the world where there exist no wild populations of these species. As an example, the GIFT strain of Nile tilapia, which is a native species in several African countries, has been disseminated to 10 countries in the Asia-Pacific including Bangladesh, P.R. China, Fiji, India, Indonesia, Malaysia, Sri Lanka, Papua New Guinea, Thailand and Vietnam (Gupta and Acosta, 2001; WorldFish Center, 2002). Since no commercial fish farming systems can be considered fully secure against the escape of some individuals, introductions of exotic species or populations for aquaculture purposes should always be considered as an introduction into the natural environment.

20.3.1 EFFECTS ON THE GENETIC DIVERSITY

Despite that exotic species are expected to be less competitive in the wild compared to native species, that have been adapted to the local environment during prolonged natural selection, there are examples of introduced species (e.g. Mozambique tilapia in Asian countries) that have been able to compete with native species and disturb ecosystems. Furthermore, changes of the habitats due to human interventions, such as agriculture, industrialization, pollution etc., might enhance the ability of populations of exotic species to be established at the expense of native species. Therefore, introductions of aquaculture species should be carefully evaluated according to the precautionary principle (see e.g. the EIFAC/ICES “Codes of practice and manual of procedures for consideration of introductions and transfers of marine and freshwater organisms”).

Changes of the habitat and/or competition with introduced species may cause natural selection in the native populations to change from stabilizing selection (i.e. there exists a balance between the environmental conditions and the average performance of the individuals) to become directional selection (i.e. cause a new genetic adaptation), resulting in changes in the genetic constitution of the populations. Furthermore, if the environmental changes are large or happen too fast, it is possible that the mortality in the native population increases as the population adapts to its new environment. As a result of this reduced effective population size, inbreeding may accumulate more rapidly, with loss of genetic variation because of genetic drift. In turn, reduced genetic variation will reduce the ability of a population to adapt to further changes in environmental conditions.

20.3.2 INTRODUCTION OF MICRO ORGANISMS AND PARASITES

The most dramatic effect of introducing non-native fish is often the unintended introduction of new and harmful microorganisms and parasites. These microorganisms and parasites may then use the local populations of fish as host organisms. As the native population is likely to be ‘naïve’ with respect to such diseases, with no development of natural defence systems, mortalities can be very high. The most common examples from Norway are the introduction of the parasite *Gyrodactylus*

salaris from Sweden and the bacterial disease furunculosis from Scotland. Both organisms followed passively when rainbow trout and Atlantic salmon were imported 20-30 years ago, and are still matters of great concern in wild populations of Atlantic salmon in Norway.

20.4 Genetic interactions between genetically improved and wild fish of the same species

The environmental impacts of genetically improved fish on wild populations of the same species will depend on the genetic differences between the genetically improved farmed fish and the wild populations, and the gene flow from farmed populations (i.e. how many genetically improved fish that succeed to breed and produce viable offspring compared with wild fish).

20.4.1 GENETIC DIFFERENCES BETWEEN FARMED AND WILD POPULATIONS

The genetic effects caused by escaped farmed fish will depend on the genetic differences between the farmed and wild populations. The impacts might be considerable if the genetic differences are large, but be insignificant if the genetic differences are small. In general, the genetic differences between farmed and wild populations of fish will depend on the origin of the farmed fish (i.e. whether they are based on local wild populations or not), which strategy that is used to genetically improve the farmed fish, the effective population size of the farmed populations, and the number of generations since the farmed populations were separated from their wild origin.

20.4.1.1 Domestication selection

The genetic constitution of farmed populations of fish will, if they are kept genetically separated from other farmed or wild populations, be changed after just a few generations. First, a directional natural selection will cause the farmed populations to adapt to the new environment in the hatcheries and fish farms, which might be considerably different from the natural environment in rivers, lakes or in the sea. If the number of broodstock used in the farmed populations is small, relatives of fish will soon start mating each other and inbreeding starts accumulating. In addition to causing inbreeding depression, accumulation of inbreeding will cause the genetic variation in the farmed populations to be reduced. The genetic difference between farmed and wild populations of fish might soon become significant, since inbreeding, genetic drift and natural selection might cause genetic traits and allele frequencies to diverge or alleles even to be fixed or lost in a closed farmed population.

20.4.1.2 Artificial selection

Selective breeding does not create new alleles, but changes the frequencies of existing alleles in favour of those with positive effects on traits that are being genetically improved. The genetic constitution of farmed populations that are genetically improved by selection will, therefore, be influenced by the origin of fish in the base population,

the traits in the breeding goal and which selection strategy that is used to change the population as defined in the breeding goal. Individual selection (mass selection) is the most commonly used selection strategy to improve growth in fish because of its simplicity. However, accumulation of inbreeding might be considerable and often cause the selection response to decrease after just a few generations due to reduced genetic variation in the breeding population. Therefore, mass selection designs have been developed to restrict inbreeding by increasing the number of parents and reducing the number of progeny tested per parent (Gjerde et al., 1996; Bentsen and Olesen, 2002).

The successful selective breeding of Atlantic salmon in Norway (Gjøen and Bentsen, 1997) and the GIFT-strain of Nile tilapia in the Philippines (Eknath and Acosta, 1998), have caused an increasing number of fish breeding programs to be established worldwide using a combined family and within-family selection strategy. The advantage with this strategy is that additional traits, such as those that are recorded on sacrificed fish (e.g. disease resistance and slaughter quality) or that have a binominal distribution (e.g. survival and age at sexual maturation), might be added to the breeding goal. Selection to improve several traits simultaneously would cause more of the genetic variation to be maintained in the breeding population and reduce the risk of allele fixation. However, a combined selection strategy requires a large number of families to be produced in each generation (i.e. ideally several hundred families) to secure high selection response in all traits without causing too much accumulation of inbreeding and loss of allelic variation.

Finally, the development of gene maps for several economically important fish species, e.g. Atlantic salmon (Lie et al., 1997) and Nile tilapia (Kocher et al., 1997) and the localization of quantitative trait loci (QTL´s), i.e. DNA sequences that are associated with the variation in quantitative traits, might facilitate marker-assisted selection. The optimal use of QTL information needs to be determined, however, especially if several complex traits (i.e. traits that are influenced by many protein-coding and regulatory genes) are to be included in the breeding goal. Simulation studies suggest that marker-assisted selection might increase the initial selection response, but reduce the potential long-term response of complex traits since the selection is too narrow to include all alleles that have positive effects (Weller, 2001). QTL selection is thus expected to result in rapid fixation of the allele(s) selected for.

20.4.1.3 Crossbreeding

Hybridisation, or crossbreeding, of individuals from different farmed populations is commonly used to counteract the negative effects of accumulated inbreeding. Since offspring receive one set of chromosomes (which contains a complete set of the DNA) from each parent, harmful effects of recessive alleles in one population might be eliminated by effects of dominant, normal alleles from another population and result in improved performance of hybrids, especially in traits related to the overall fitness of individuals (i.e. cause heterosis). Highly inbred lines of fish for hybridisation schemes might also be produced experimentally using gynogenesis (i.e. self-fertilisation); a strategy that is more commonly used when developing genetically improved varieties of plants.

However, the heterosis effects cannot be transferred from one generation to the other since it depends on the combinations of gene pairs in an individual (i.e. its complete genotype). Therefore, crossbreeding schemes require that two or more breeding populations are maintained in separate rearing units, or even better, are continuously improved by selection within each population. Although it is possible to develop hybrid offspring that are very heterozygous (i.e. they have received different alleles in most gene pairs), the history of developing and maintaining the parents' populations might cause these farmed hybrid populations to become genetically much different compared to their wild origin.

20.4.1.4 Gene transfer (GMO)

Several techniques are available to introduce gene constructs (i.e. new combinations of protein-coding and regulatory DNA sequences, which might originate from the same or other species) into the fertilized eggs of fish, such as microinjection, electroporation and particle bombardment. Gene transfer in fish has currently been restricted to the transfer of growth hormone and anti-freeze genes (Hew et al., 1995). The expression of these genes are increased substantially by selecting appropriate promoters, and the increased growth rate realized in some transgenic fish is primarily caused by the tissue-specific promoter directing growth hormone synthesis to occur in the liver in addition to the pituitary gland.

However, the transferred gene constructs are properly incorporated in less than 10% of the micro-injected eggs, and each transgenic animal is unique since they might receive different number of gene constructs located at different sites on their chromosomes (Hackett, 1993). Therefore, the development of farmed populations of transgenic fish might require several generations of selection to multiply the number of individuals carrying the transferred gene constructs. The concern of farming transgenic fish (often referred to as Genetically Modified Organisms, GMO´s) is that escapees, in contrast to those from selective breeding programs, might cause new gene constructs to be disseminated in wild populations of fish.

20.4.2 GENE FLOW FROM GENETICALLY IMPROVED BREEDS TO WILD POPULATIONS

The gene flow from genetically improved breeds to wild populations will depend on the number of escapees, their breeding success and the survival rate of their offspring. No commercial aquaculture systems can be considered fully secure against the escape of farmed fish. However, proper farm constructions (i.e. secured inlets and outlets of fish ponds, strengthened or doubled layer of nets in fish cages etc.) and good farm management may reduce the number of escapees. Furthermore, the breeding success of escaped farmed fish and the survival rate of their offspring might be affected by which technology that is used to genetically improve the fish.

20.4.2.1 Sterile individuals

Chromosome manipulations (e.g. by temperature shock of fertilized eggs) can be used to produce triploid individuals of fish (i.e. individuals with three sets of chromosomes).

Although it is questionable whether triploid fish have superior growth compared to their diploid counterparts (review by Tave, 1986a), it might be a suitable technique to produce sterile individuals and restrict the environmental impact of escapees. Some triploid fish might develop sexual characteristics, but they will not succeed in reproducing due to their extra set of chromosomes. However, large numbers of sterile escapees may have other impacts on wild stocks in the environment, e.g. through competitive behaviour and feeding.

20.4.2.2 Mono-sex populations

It might be beneficial to farm mono-sex populations of fish if one of the sexes grow faster than the other (e.g. males of tilapia or females of Atlantic halibut), if one of the sexes becomes sexually mature too early (e.g. males of Atlantic salmon) or if there exists a problem with uncontrolled breeding in the farming environment (e.g. tilapia). Several techniques, such as chromosome manipulations (e.g. to produce YY-males of tilapia, Beardmore et al., 2001), gynogenesis (e.g. to produce all-female Atlantic salmon, Thorgaard et al., 2002) and inter-specific hybridisation (e.g. Nile tilapia x Blue tilapia), are commonly used to produce mono-sex populations. Although mono-sex populations might reduce the problem with uncontrolled breeding in fish farms, they cannot avoid a gene flow from farmed to wild populations where both sexes are available. The same number of escapees from mono-sex populations might even cause a larger gene flow to wild populations than escapees from mixed populations, since they only have to compete with wild individuals of the same sex.

20.5 Case study: Interactions between farmed and wild Atlantic salmon in Norway

Ever since the early start of the farming of Atlantic salmon in Norway more than 30 years ago, farmed salmon have escaped from the farm cages and been observed together with wild Atlantic salmon in Norwegian rivers, where they may spawn (Lura and Økland, 1994). The number of escaped farmed salmon increased dramatically at the end of the 1980´s, and it has been estimated that about 1.6 millions salmon escaped annually in the period 1988-92. Although, the annual number of escapees was reduced to between 250,000 – 500,000 in the 1990´s, the number has again increased during the last couple of years. The total number of wild smolts migrating from Norwegian rivers has been estimated to be about 6 millions, and the annual number of sexually mature salmon returning to the same rivers has varied between 200,000 – 300,000 individuals in the 1980´s and 1990´s. Since farmed salmon might constitute a large proportion of the spawners in some rivers, there is a considerable public concern in Norway concerning the possible environmental impacts of escaped farmed salmon.

Farmed Atlantic salmon in Norway may not be a completely representative case study of genetic interactions between farmed and wild populations of fish. Unlike most other farmed stocks, the farmed populations of Norwegian salmon are entirely based on genetic material collected from local wild populations, and no known introductions have been made from populations outside the country. In addition, the number of parallel farmed populations that are under selection is much larger than in most other

aquaculture species, and random genetic changes caused by genetic drift is likely to differ from one population to another.

20.5.1. GENETIC COMPOSITION OF NORWEGIAN FARMED SALMON

Since all parents of the Norwegian aquaculture founder populations were caught in Norwegian rivers, farmed salmon in Norway cannot carry alleles that do not occur naturally in Norwegian wild salmon. As previously described in Chapters 16 and17, the best-documented selective breeding program with Atlantic salmon in Norway is a family based program initiated by AKVAFORSK at the beginning of the 1970´s (Gjedrem et al., 1991; Gjøen and Bentsen, 1997). Because farmed Atlantic salmon in Norway has a 4-year generation interval, four genetically distinct populations were established to provide the salmon farmers with genetically improved smolts each year. During the years 1971-74, genetic materials (i.e. milt and eggs) were collected from a total of 190 males and 430 females from more than 40 Norwegian rivers. In addition, several other breeding programs based on mass selection were started in the 1970-ies and early 1980-ies with sub-populations of similar origins.

A second family based breeding program (NFA) was established in 1985 with another four populations to further secure the supply of genetically improved Atlantic salmon. The station received representative samples from all families produced at AKVAFORSK during the years 1986-89 and some additional fish from other breeding populations in Norway. To reduce the accumulation of inbreeding and fixation/loss of alleles, eyed-eggs from 10-20 families have been exchanged between the two parallel breeding programs each year. In 1992, the two parallel breeding programs (at AKVAFORSK and NFA) were organized as a common breeding company (AquaGen).

Recently, four additional breeding populations have been established by a new breeding company (SalmoBreed) under the supervision of AKVAFORSK Genetics Center, based on available material of farmed Atlantic salmon in Norway.

20.5.1.1 Genetic changes of the farmed populations

Today, both Aqua Gen and SalmoBreed are organized with a breeding station, several test stations (i.e. commercial farms and fish labs), broodfish stations and multiplier units (i.e. commercial hatcheries) to disseminate the genetically improved Atlantic salmon to the on-growing farms in Norway with almost no time delay. Therefore, the genetic performance and the allelic variation of most farmed Atlantic salmon in Norway is similar to that of the pooled breeding populations developed by the two breeding companies. Both breeding companies are using combined family and within-family selection to genetically improve several economically important traits as defined in the breeding goal. The number of traits in the breeding goal has been gradually increased as methodologies have been developed to record new traits: growth recorded as body weight at harvest (1976), age at sexual maturation (1983), resistances against furunculosis and infectious salmon anaemia (1991), filet colour (1994) and total filet fat and fat distribution (1994). The accumulated genetic gain after five generations of selection were estimated to be 85% increased body weight at harvest, 12.5 %-units reduction of early sexually maturing fish (Gjerde and Korsvoll, 1999), and 20%

improved feed efficiency as a correlated response of faster growth (Thodesen et al., 1999). The broad breeding goal is expected to result in a rather low selection pressure on each of a large number of alleles.

Initially, the AKVAFORSK/AquaGen breeding programs produced and tested 120 families per year. However, the number of families has been increased substantially as new technology for tagging (i.e. PIT-tags) became available. Today, the family based breeding programs are producing and testing about 300 families per year. The effective population size in the breeding populations under the initial family structure was estimated to range between 33-125 individuals, which is similar to the international recommendations for short-term conservation of genetic variation in animal populations (i.e. 50 individuals). Furthermore, the total effective population size across all breeding populations of Atlantic salmon in Norway is similar to that recommended for long-term conservation of animal populations, i.e. 500-5000 individuals (Frankel and Soulé, 1981; Lande, 1995).

However, molecular genetic studies have indicated a reduced heterozygosity within single breeding populations of farmed salmon (observed on 5th and 6th generations salmon in the breeding program at Aqua Gen) when compared with river populations of wild Atlantic salmon in Norway (Skaala et al., 1999). Furthermore, differences in allele frequencies between farmed populations and some of the original river populations have also increased (Mjølnerød et al., 1997). Still, the heterozygosity across all breeding populations has not been significantly reduced, suggesting that the loss of allelic variation differs from one breeding population to another. Estimated allele frequencies of 15 enzyme-coding genes suggest that about half of the alleles in wild Atlantic salmon have a frequency less than 1% and an additional fifth of the alleles have frequencies less than 10% (Anon., 1999). Based on an effective population size of 30-100 individuals, it may be estimated that the average time before neutral alleles with an initial frequency of 1% are lost because of genetic drift is 6-19 generations, and for alleles with an initial frequency of 10% the average time is 31-102 generations (calculated according to Kimura, 1970). Therefore, even if genetic drift differs randomly from one breeding population to another, it is likely that some alleles that are rare in the wild salmon populations have been lost in the entire farmed salmon population in Norway after 7-8 generations of selection.

20.5.2 GENETIC COMPOSITION OF THE WILD RIVER POPULATIONS

Since anadromous Atlantic salmon usually return to the same locality as they were born in, genetically differentiated populations might evolve as a result of natural selection and genetic drift. It is, often assumed therefore that each of 450 Norwegian salmon rivers or localities may have developed genetically unique populations of Atlantic salmon, and that each population should be conserved separately according to the previously cited stock management recommendations.

However, official catch statistics from Norwegian rivers suggest that most of these rivers have very small broodstock populations of Atlantic salmon, while a few rivers have very large populations. Recent catch statistics are considered to be quite reliable,

and the reported catches in the rivers during 1984-1986 are shown in Figure 20.1. The overall catch rate in Norwegian rivers is about 50%. The number of spawners in each river during the spawning season will consequently be of the same magnitude as the number of salmon caught during the fishing season. This means that more than 75% of the river populations numbers less than 50 spawners per year and more than 90% of the populations numbers less than 500 spawners.

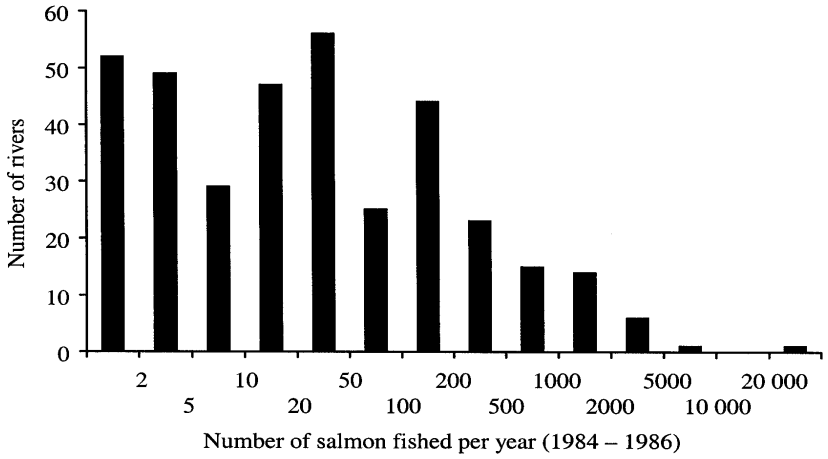


Figure 20.1. The number of adult salmon caught per year in each of 362 Norwegian rivers where catches were reported in 1984-1986, according to the official catch statistics. The average catch rate is about 50%, leaving about the same number of salmon in the rivers to spawn. Note the non-linear scale on the horizontal axis.

Furthermore, studies suggest that the effective population size may be about one fifth of the number of fish spawning in the river (Altukhov and Salmenkova, 1994). If the Norwegian river populations of Atlantic salmon were really genetically separated from each other, very few of them would then be large enough to avoid accumulation of inbreeding depression and loss of genetic variation (because of genetic drift) at a rate that would threaten their existence within a limited number of generations. Consequently, the persistence of most of the small river populations seems to depend on a regular migration of spawners between the rivers. Such sub-divided populations are often referred to as meta-populations. Studies suggest that on average 4% (range 0-20%) of the sexually mature salmon return to other rivers than the one they were hatched in (Stabell, 1984). Lost alleles might then be reintroduced and inbreeding counteracted across the meta-population. The relatively few river populations of Atlantic salmon in Norway that is large enough to maintain their viability and genetic variation over time, may then serve as reservoirs of genetic variation that is regularly disseminated to the smaller populations by migrating fish.

This implies an effective migration of spawners between river populations that will counteract a possible genetic diversification of the populations caused by genetic drift or diverse natural selection in the different river environments. The magnitude of

random genetic diversification may be studied by analysing differences in the frequencies of neutral alleles between populations. Estimates based on the frequency of nine variable enzyme-coding genes in several Norwegian river populations, indicate that 93.5% of the heterozygosity was caused by variation within river populations and only 6.5% was caused by differences between river populations (Ståhl and Hindar, 1988), a result that has been confirmed by several later studies. If this is also true for neutral polygenic traits, the average random diversification between river populations may be illustrated as shown in Figure 20.2 (a), i.e. the populations will be largely overlapping. Comparing this figure to Figure 4.4, it is evident that the genetic differences between populations in terms of allelic composition will be small. Although the natural gene flow caused by migrating spawners may be restricted, it nevertheless seems to be sufficient to prevent Norwegian river populations from diverging much because of genetic drift.

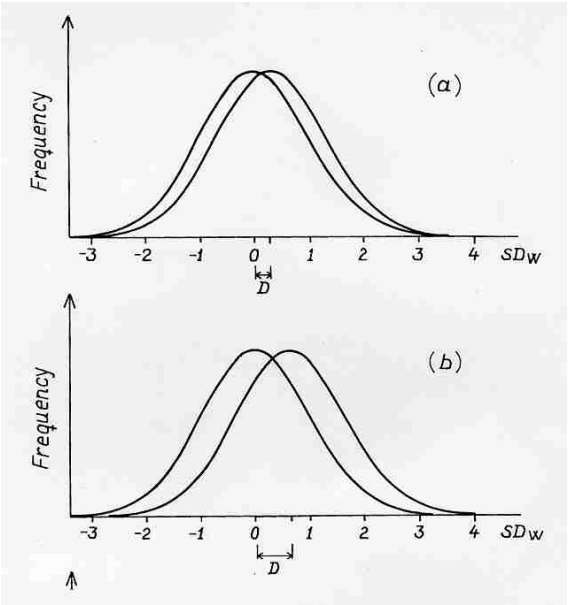


Figure 20.2. Average genetic distribution of a polygenic trait in two populations when 5% (a) or 25% (b) of the total genetic variation is variation between populations.

Still, it is an established fact that Norwegian river populations may express differences in average values of several complex, polygenic traits such as body weight, age at smoltification, age at sexual maturation and the timing of smolt migration, return of spawners to the rivers, breeding, hatching etc. (Huitfeldt-Kaas, 1946; Jonsson et al., 1991; Heggberget et al., 1993). These are all phenotypic observations that may be strongly influenced by systematic environmental effects in the different rivers. But even if most river populations of Atlantic salmon in Norway are probably too small to accumulate specific, long-term evolutionary genetic adaptation to their local environments, the extremely strong selection in wild salmon (only about 0.001% of the

eggs will give rise to a successful adult spawner) may produce large genetic fluctuations from one generation to the next. This short-term response to selection will be in balance with the homogenizing effects of migration, and thus generate non-evolutionary genetic differentiation between populations. In the larger populations where the response to natural selection is less influenced by migrating fish from other populations and by genetic drift, long-term genetic adaptations may accumulate if the divergent environmental conditions are historically stable.

Studies of Atlantic salmon from 40 different Norwegian rivers, carried out in standardised tank or cage environments, indicate that 2-8% of the observed phenotypic variation of fitness traits such as hatchability of eggs, survival rate and growth of fry, age at smoltification and growth rate at various stages was due to differences between river populations (Kanis et al., 1976; Refstie et al., 1977; Refstie and Steine, 1978; Gunnes and Gjedrem, 1978). Since the heritability of these traits was between 0.1 and 0.3, it may be estimated that 10-20% of the total variation in genetic values for these traits was associated with differences between river populations. This is a significantly larger differentiation than the random differences referred to above, indicating that differential selection is indeed effective in Norwegian river populations. Figure 20.2 (b) shows the average genetic differentiation between two populations for a polygenic trait when 25% of the total variation is caused by differences between populations. Again, when comparing to Figure 4.4, this is still a rather weak differentiation that may be achieved without large differences in allele frequencies if the trait is affected by many genes.

20.5.3 GENE FLOW FROM FARMED TO WILD POPULATIONS

The gene flow from farmed to wild populations of Atlantic salmon can be estimated by recording the proportion of farmed salmon caught in the fisheries at different locations and life stages of the fish, from studies where tagged farmed salmon of different ages have been released at different times of the year to follow the behaviour and the survival of escapees, and by comparing the relative fitness and reproductive performance of farmed and wild salmon in natural environments. If each river population is considered as a genetically unique product of evolution, the gene flow to each population should be evaluated separately. On the other hand, if the entire Norwegian population of Atlantic salmon is regarded as one meta-population, the entire gene flow may be considered as a whole, as in the following discussion.

20.5.3.1 Representation of escaped farmed salmon in wild populations

The proportion of farmed salmon caught in the fisheries varies at different locations (Figure 20.3). About 40% of the salmon caught in the sea fisheries near the Faeroe Islands in the early 1990's was classified as farm escapees, but the proportion of escapees has been reduced to about 25% in later years. In the fisheries along the Norwegian coastline, 37-54% of the catch was classified as escaped farmed salmon in the period 1989-96. In the Norwegian fjords, where most of the salmon are on their way to breed in the rivers, it was estimated that 10-21% of the salmon catch originated from fish farms in the same period. Still, only 4-7% of the salmon caught in the Norwegian rivers during the fishing season were classified as escapees. However, escaped farmed

salmon seem to have a delayed spawning migration compared with wild salmon, and it has been estimated based on studies in some rivers after the end of the fishing season that on average 17% of the Atlantic salmon broodstock in Norwegian rivers originate from fish farms (Anon., 1999). The representation of escaped salmon was highest in the small rivers and close to zero in some of the biggest rivers.

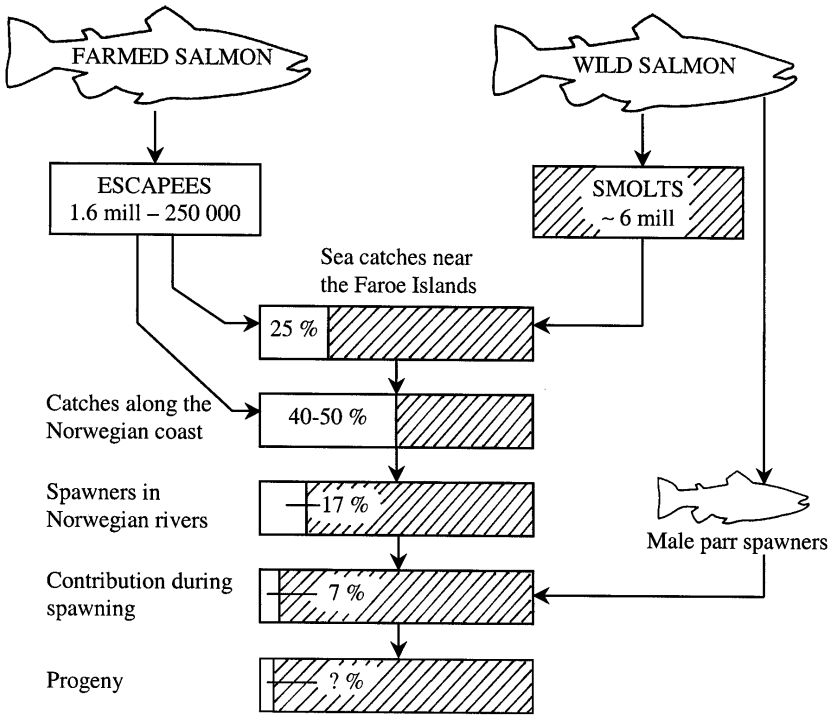


Figure 20.3. Gene flow from escaped farmed Atlantic salmon in Norway to the pooled population of Norwegian river stocks.

20.5.3.2 Relative breeding success of escaped farmed salmon

The gene flow from farmed populations of Atlantic salmon is also influenced by the relative breeding success of the escapees compared with wild salmon. In a large tank experiment, farmed male and female breeders originating from the third generation of the selective breeding program at AKVAFORSK (now managed by AquaGen) had only 3% and 30% breeding success, respectively, as compared to their wild broodstock competitors (Fleming et al., 1996). However, the relative breeding success of farmed salmon might be higher if they escape at early life stages. In an experiment where hatchery-produced Atlantic salmon of wild origin were released as smolts in their parents' river (Fleming et al., 1997), returning males and females had 51% and 100% breeding success, respectively, compared to wild individuals that were hatched naturally in the same river. The breeding success would probably be lower than this if the

hatchery smolt had been taken from a domesticated population. Considering the estimates that about half of the escapees are escaping at early stages (Lura and Økland, 1994) and that wild parr males sometimes might breed without migrating to seawater, the total breeding success of escaped farmed Atlantic salmon may then be estimated to be maximum 40% compared with that of wild salmon (Figure 20.3).

20.5.3.3 Survival of offspring

Combining the estimated representation and breeding success of escaped farmed salmon, suggests that the gene flow from farmed to wild populations of Atlantic salmon in Norway at the stage of spawning is maximum 6-7% (Anon., 1999). The gene flow will, however, further depend on the relative survival rates of the offspring of farmed and wild salmon. Several Norwegian tank tests (Hindar et al., 1991) and an Irish field test (McGinnity et al., 2003) suggest that offspring of farmed salmon grow faster and have a more competitive behaviour compared with offspring of wild salmon. However, the survival of the offspring of farmed salmon was much lower in all stages of life in the Irish study. The gene flow from farmed to wild populations of Atlantic salmon in Norway is, therefore, expected to be lower than the estimate given above.

20.5.4 EFFECTS OF GENETIC INTERACTIONS BETWEEN FARMED AND WILD ATLANTIC SALMON

The effects of genetic interactions between farmed and wild Atlantic salmon in Norway will in the following be discussed based on expected changes of the allelic variation, future expression of fitness traits and the productivity of wild river populations.

20.5.4.1 Effects on the allelic constitution of wild populations

Since salmon farming in Norway is entirely based on genetically improved fish originating from Norwegian rivers, farmed escapees cannot introduce other alleles into local river populations than those that already exists or might be introduced by migrating wild salmon from other rivers. However, farmed populations of Atlantic salmon may have different allele frequencies compared with the wild populations because of the ongoing selective breeding, and the genetic differences will increase as a result of genetic drift as long as the farmed populations are kept genetically isolated. Since the escaped farmed salmon represent a one-way gene flow from the farmed to the wild populations, it is expected that the frequencies of neutral alleles in the wild populations gradually will change and become similar to the average of all breeding populations. Several studies have shown that frequencies of neutral alleles of wild populations have been changed because of escaped farmed salmon both in North Ireland (Crozier, 1993), Ireland (Clifford et al., 1998a, b) and in Norway (Skaala and Hindar, 1998). Therefore, it is expected that neutral alleles that are lost in all breeding populations might also, in a long perspective, be lost in the wild populations. However, frequencies of other alleles, which affect the adaptation of wild salmon in its present environment (i.e. fitness traits), are less sensitive to changes since they tend to be stabilized by natural selection. The risk of loosing fitness alleles in a population is, therefore, far less than loosing neutral alleles. The effects on the genetic constitution of wild populations might be reversible if the number of escaped farmed salmon is reduced and migrating fish from other river populations reintroduce lost alleles. However, it is

not expected that the natural selection in these rivers will retrieve (or maintain) exactly the same allele frequencies, since the optimal adaptation of an individual might be the final result of many different combinations of allele frequencies.

20.5.4.2 Effects on genetic adaptation of wild populations

Most fitness traits are affected by many protein-coding and regulatory genes. As a result, the genotypic value of these traits will not be determined by the frequency of single alleles, and changes of frequencies or loss of some alleles might be compensated by changes in the frequencies of alleles in other loci. Therefore, the possible reduction of allelic variation because of escaped farmed fish has to be of a certain magnitude before it reduces the ability of wild populations to respond to the ongoing natural selection. The frequency of genetically adapted offspring is expected to be much higher in locally adapted river populations than among those produced by escaped farmed fish. This does not exclude, however, that some offspring of farmed salmon, and especially hybrids between farmed and wild fish, also might inherit well-adapted fitness traits. If this happens, these individuals are not likely to represent a severe threat to future genetic adaptation of the local river populations, but might be expected to have similar effects to those caused by natural migration between wild river populations. However, the equilibrium between genetic effects of migration and natural selection may not always be optimal for the natural adaptation of river populations to their local environments (referred to as a genetic load). If the migration is reduced, the genetic load will also be reduced. Therefore, it is still important to strengthen all efforts to restrict the number of escaped farmed salmon.

Since some large river populations of Atlantic salmon might serve as reservoirs for regular introductions of new genetic variation of fitness traits through migration to smaller river populations that are more severely affected by farmed escapees, the Norwegian fisheries management authorities have proposed to establish special areas where fish farming should be restricted to protect these important river strains.

20.5.4.3 Effects on the productivity of wild populations

Farmed salmon will, most likely, become less adapted to the natural habitats in the rivers as the domestication in the selective breeding programs continues. Even if this will reduce the gene flow from escaped farmed salmon into the wild populations (as the survival rate of farmed salmon in the wild is reduced), escaped farmed salmon might still affect the productivity of wild salmon in the rivers. Observations suggest that escaped farmed salmon may have a delayed spawning migration into the rivers, and that they might cause problems by digging up fertilized eggs already spawned by the wild salmon (Lura and Sægrov, 1991). In such cases, the offspring of farmed salmon will most likely hatch too late to survive, while the production of wild offspring will be reduced. Furthermore, Norwegian and Irish experiments (Hindar et al., 1991; McGinnity et al., 2003) suggest that offspring of farmed salmon may have a more dominant behaviour compared with offspring of wild salmon, and consequently occupy scarce feed resources and habitats during early life stages. This may reduce the production of more viable wild smolt in the river. Since the production of viable, migrating smolt might be essential for the number of sexually mature salmon returning back to the rivers, the productivity of the wild river populations may then be reduced.

20.6 Summary

Aquaculture is often based on the culture of introduced species or of exotic or genetically improved populations of indigenous species. Since no commercial aquaculture systems may be considered fully safe against unintended releases of the farmed organisms, the establishment of aquaculture operations should always be regarded as an introduction of the organisms into the surrounding environment. In most cases, introduced species are expected to compete poorly compared with native species that are genetically adapted to their environment. However, there are many examples of harmful effects of introductions on the local genetic diversity, and international recommendations and codes of practice for transfer of organisms should be followed. A special attention should be paid to the possible introduction of disease organisms that may follow the aquaculture species. The culture of exotic or genetically improved populations of indigenous populations may represent an additional problem, since escapees may interbreed with wild broodstock. In cases where local wild stocks have a long history of evolutionary adaptation to a particular environment, interbreeding may disturb the adaptation in the hybrid offspring. The most thoroughly investigated example of genetic interactions between farmed and wild populations of fish is the Atlantic salmon case in Norway. Results from several studies are presented and discussed. The case may not be completely representative, since the farmed salmon in Norway originates from a number of more or less independent breeding programs, all with an exclusive genetic origin from Norwegian river stocks.

APPENDIX A

Portion selected (p), number of standard deviation from population average to point of truncation (x), selection intensity (i), for a normally distributed trait.

p%	x	i	p%	x	i
0.01	3.719	3.960	09	1.341	1.804
0.05	3.291	3.554	10	1.282	1.755
0.10	3.090	3.367	11	1.227	1.709
0.20	2.878	3.170	12	1.175	1.667
0.30	2.748	3.050	13	1.126	1.627
0.40	2.652	2.962	14	1.080	1.590
0.50	2.576	2.892	15	1.036	1.554
0.60	2.512	2.834	16	0.994	1.521
0.70	2.457	2.784	17	0.954	1.489
0.80	2.409	2.740	18	0.915	1.458
0.90	2.366	2.701	19	0.878	1.428
1.00	2.326	2.665	20	0.842	1.400
2.00	2.054	2.421	25	0.674	1.271
3.00	1.881	2.268	30	0.524	1.159
4.00	1.751	2.154	35	0.385	1.058
5.00	1.645	2.063	40	0.253	0.966
6.00	1.555	1.985	45	0.126	0.880
7.00	1.476	1.918	50	0.000	0.798
8.00	1.405	1.858	60	-0.25	0.644

REFERENCES

Allen, S.K. 1983. Flow cytometry. Assaying experimental polyploid fish and shellfish. Aquaculture, 33: 317-328.

Allen, Jr. S.K. 1998. Commercial applications of bivalve genetics: not a solo effort. World Aquaculture, 29: 38-43

Allen, Jr. S. K., and Bushek, D. 1992. Large-scale production of triploid oysters, *CrassostreaVirginica* (Gmelin), using "stripped" gametes. *Aquaculture,* 103: 241-251.

Allen, S.K., Gagnon, P.S. and Hidu, H. 1982. Induced triploidy in the soft-shell clam. Cytogenetic and allozymic confirmation. The Journal of Heredity, 73: 421-428.

Allen, S.K. and Stanley, J.G. 1978. Reproductive sterility in polyploid brook trout, *Salvelinus fontinalis*.Trans. Am. Fish. Soc. 107: 473-478.

Allen, S.K. and Stanley, J.G. 1979. Polyploid mosaics induced by cytochalasin B in landlocked Atlantic salmon, *Salmo salar*. Trans. Am. Fish. Soc. 108: 462-466.

Allen, S.K. and Stanley, J.G. 1981. Mosaic polyploid Atlantic salmon *(Salmo salar*) induced by cytochalasin B. Rapp. P.-v Reun. Cons. perm. int. Explor. Mer. 178: 509-510.

Altukhov, Y.P. and Salmenkova, E. A. 1994. Straying intensity and genetic differentiation in salmon populations. Aquacult. Fish. Manage. 25 (Suppl. 2), 99-120.

Andersen, B.B. 1977. Genetic studies concerning growth rate in cattle, body development and feed efficiency (Genetiske undersøkelser vedrørende kvaegets tilvaekst, kropsudvikling og foderudnyttelse). Rep-No. 488, Natl. Inst. Inst. Anim. Sci., Copenhagen, 137 pp.

Anon, 1999. Til laks åt alle kan ingen gjera? NOU 1999:9.

Anon, 2004 Trout news, No 37, January 2004, CEFAS and defra: 5-8.

Aulstad, D., Gjedrem, T. and Skjervold, H. 1972. Genetic and environmental sources of variation in length and weight of rainbow trout *(Salmo gairdneri).* J. Fish. Res. Bd. Canada, 29: 237-241.

Aulstad, D and Kittelsen, A. 1971. Abnormal body curvatures of rainbow trout *(Salmo gairdneri)* inbred fry. J. Fish. Res. Bd. Canada, 28: 1918-1920.

Ayles, G.B. and Baker, R.F. 1983. Genetic differences in growth and survival between strains and hybrids of rainbow trout *(Salmo gairdneri)* stocked in aquaculture lakes in the Canadian prairies. Aquaculture, 33: 269-280.

Bakken, M., Vangen, O. and Rauw, W. 1998. Biological limits to selection and animal welfare. Invited paper, Proc. 6th World Congress of Genetics Applied to Livestock Production, 27: 381-389.

Bakos,J. 1979. Crossbreeding Hungarian races of common carp to develop more productive hybrids. In: T.V.R. Pillay and W.A. Dill (Editors), Advances in Aquaculture. Fishing News Books Ltd, Farnham, Surrey, 633-635 pp..

Bakos, J. 1987. Selective breeding and intraspecific hybridisation of warm water fishes. Proc. World Symposium on Selection, Hybridisation, and genetic Engineering in Aquaculture, Bordeaux, I: 303-311.

Barber, B., Davis, C. and Hawes, R. 1998. Genetic improvement of oysters for the Maine quaculture industry. Abstract of the first Annual Northwest Aquaculture Conference nd Exposition, Nov. 1998: 35.

Bardach, J.E., Ryther, J.H. and McLarney, W.O. 1972. Aquaculture. The farming and usbandry of fresh water and marine organisms. John Wiley and Sons. New York, hichester, Brisbane, Toronto.

Barlow, R. 1983. Benefit-cost analyses of genetic improvement program for sheep, beef cattle and pigs in Ireland. Ph. D. Thesis, University of Dublin. Ref. by E. P. Cunninham, Present and future perspective in animal breeding research. XV. International Congress of Genetics, New Dehli, India, 12-21. December, 1983, 19 pp.

Beacham, T.D. and Evelyn, T.P.T. 1992. Genetic variation in disease resistance and growth in chinook, coho and chum salmon with respect to vibriosis, furunculosis, and bacterial kidney disease. Trans. Amer. Fish. Soc., 121: 456-485.

Beardmore, J.A., Mair, G.C. and Lewis, R.I. 2001. Monosex male production in finfish as exemplified by tilapia: application, problems, and prospects. Aquaculture, 197: 283-301.

Beilharz, R., Luxford, B.G. and Wilkinson, J.L. 1993. Quantitative genetics and evolution: Is our uderstanding of genetics sufficient to explain evolution? J.Anim.Breed.Genet. 110:161-170.

Bencze, A.M. 1988. Flesh quality of triploid rainbow trout (Kvalitet på triploid regnbue ørret). Msc. thesis, Agricultural University of Norway, In Norwegian, 69 pp.

Bennet, R.M. 1996. People's willingness to pay for farm animal welfare. Animal Welfare, 3-11.

Bentsen, H.B. 1990. Application of breeding and selection theory on farmed fish. Proceedings of the 4th World Congress on Genetics Applied to Livestock Production, Edinburgh 23-27 July, 1990. 149-158.

Bentsen, H.B. 1994. Genetic effects of selection on polygenic traits with examples from Atlantic salmon. S*almo salar* L. Aquaculture and Fisheries Management, 25:89-102.

Bentsen, H.B., Eknath, A.E., Palada-deVera, M., Danting, J.C., Bolivar, H.L., Reyes, R.A., Dionisio, E.E., Longalong, F.M., Circa, A.V., Tayamen, M.M., and Gjerde, B. 1998. Genetic improvement of farmed tilapias: Growth performance in a complete diallel cross experiment with eight strains of *Orechromis niloticus*. Aquaculture, 160:145-173.

Bentsen, H. B., Eknath, A.E., Rye, M., Thodesen, J. and Gjerde, B. 2004. Genetic improvement of farmed tilapias: response during five generation of selection for body weight at harvest in *Oreochromis niloticus.* In manuscript.

Bentsen, H.B., Gjedrem, T. and Hao, N.V. 1996a. Breeding plan for silver barb *(Puntius Gonionotus)* in Vietnam: Improve growth rate. INGA, July 1996, ICLARM, Report No.3.12 pp

Bentsen, H.B., Gjedrem, T., Tien, T.M. and Dan, N.C. 1996b. Breeding plan for Nile tilapia in Vietnam: Combined multitrait selection. INGA, July 1996, ICLARM, Report No. 1. 10 pp.

Bentsen, H.B. and Gjerde, B. 1994. Design of fish breeding programs. Proceedings of the 5th World Congress on Genetics Applied to Livestock Production, Selection and quantitative genetics; Growth; Reproduction; Lactation; Fish; Fiber; Meat, Guelph August 7-12, 19: 353-359.

Bentsen, H,B. and Olesen, I. 2002. Designing aquaculture mass selection programs to avoid high inbreeding rates. Aquaculture, 204: 349-359.

Benzie, J.A.H., Kenway, M., Ballment, E.,Fresher, S. and Trott, L. 1995. Interspecific hybridisation of the tiger prawns Penaeus monodon and penaeus esclentus. Aquaculture, 133:103-111.

Benzie, J.A.H., Kenway, M. and Trott, L. 1997. Estimates for the heritability of size in juvenile Penaeus monodon prawns from half-sib matings. Aquaculture, 152: 49-53.

Berg, P. and Henryon, M. 1998. A comparison of mating designs for inference on genetic parameters in fish. In: Proc. of the 6th World Congress on Genetics Applied to Livestock Production, Armidale, NSW, Australia, 27:115-118

Betsch, D. F. 1994. Principles of biotechnology. Biotechnology Information Series (Bio-1) North Central Regional Extension Publication Iowa State University - University Extension.

Billard, R. 1978. Changes in structure and fertilizing ability of marine and freshwater fish spermatzoa diluted in media of various salinities, Aquaculture 14: 187-198.

Billingsley, L.W. 1981 (ed.). Proceedings of the stock concept international symposium. Can. J. Fish. Aquat. Sci. 38:1457-1921.

Blaxter, J.H.S. 1953. Sperm storage and crossfertilization of spring and autumn spawning herring. Nature 172: 1189-1190.

Bolivar, R. 1999. Estimation of response to within-family selection for growth in Nile tilapia. Diss. Abst. Int. Pt. B-Sci and Eng. 60(3): p.934.

Bolla, S. and Refstie, T. 1985. Effect of cytochalasin B on eggs of Atlantic salmon and rainbow trout. Acta Zoologica (Stockh.), 66 (3): 181-188.

Bondari, K. 1983. Response to bidirectional selection for body weight in channel catfish. Aquaculture 33:73-81.

Bondari, K. 1984. A study of abnormal characteristics of channel catfish and blue tilapia. Proc. Ann. Ci\onf. Southeast. Assoc. Fish Wild. Agen. 35 (1981): 566-578.

Bondari, K. and Dunham, R.A. 1987. Effects of inbreeding on economic traits of channel catfish. Theor.Appl.Genet. 74: 1-9.

Bongers, A.B.J. 1997. Development and application of genetically uniform strains of common carp *(Cyprinus carpio L.)*. Doctoral thesis, Wageningen Agricultural University, Wageningen, the Netherlands, 155 pp.

Braden, J.B. and Kolstad, C.D. 1991. Measuring the demand for environmental quality. Elsevier Science Publishers, Amsterdam, 370 pp.

Brichette, I., Reyero, M.I. and Garcia, C. 2001. A genetic analysis of intraspecific competetion for growth in mussel culture. Aquaculture, 192: 155-169.

Bridges, W.R. and von Limbach, B. 1972. Inheritance of albinism in rainbow trout. J. Hered. 63: 152-153.

Brier et al. 1940 pages 153-160, in Proc. Amer. Soc. An. Prod.

Brooker, A.L., Cook, D., Bentzen, P., Wright, J.M. and Doyle, R.W. 1994. Organization of microsatellites differs between mammals and cold-water teleost fishes. Canadian Journal of Fisheries and Aquatic Sciences, 51: 1959-1966.

Brown, C., Woolliams, J. and McAndrew, B. 2004. Factors influencing effective population size in commercial populations of gilthead seabream, *Sparus aurata*. Aquaculture, in press

Brusle, S. 1987. Sex-inversion of the hermaphroditic protogynous teleost *Coris julis L. (Labridae).* Journal of Fish Biology. 30: 605-616.

Brännäs, E. and Alanärä, A. 1992. Feeding behaviour of the Arctic charr in comparison with rainbow trout. Aquaculture 105, 53-59.

Bulmer, M.G. 1985. The matematical theory of quantitative genetics. Clarendon Press, Oxford.

Bulmer, M.G. 1971. The effect of selection on genetic variability. Am. Nat. 105: 201-21.

Buyukhatipoglu, S and Holtz, W. 1978. Preservation of trout sperm in liquid or frozen state. Aquaculture 14: 49-56.

Cabellero, A., Santiago, E. and Toro, M. 1996. Systems of mating to reduce inbreeding in selected populations. Anim.Sci. 62: 431-442.

Cameron, N.D. 1997. Selection indices and prediction of genetic merit in animals breeding. CAB International, Wallingford, Oxon OX10 8DE, UK. ISBN 0 85199 196 6.

Carlborg, O., Kerje, S., Schutz, K., Jacobsson, L., Jensen, P. and Andersson, L. 2003. A global search reveals epistatic interaction between QTL for early growth in the chicken. Genome Res. 13(3): 413-21.

Carlin, B. 1969. Salmon tagging experiments. Laksforskningsinstitutet (Swedish Salmon Research Institute) Medd. 2-4: 8-13.

Casella, G. and George, E.I. 1992. Explaining the Gibbs sampler. The American Statistician, 46: 167-174.

Chapman, A.B. 1962. Biometrical genetics related to breeding plans. Lecture notes. Department of Genetics, University of Wisconsin, USA.

Chen, F.Y. 1969. Preliminary studies on the sex-determinating mechanism of Tilapia mosambica Petersand T. hornorum Trewawas. Verh. Int. Ver. Theor. Angew. Limnol., 17: 719-724.

Cerda, C., Neira, R. and Barria, N. 2003. Genetic and phenotypic parameters for harvest weight, color flesh and fat content in coho salmon (Oncorhynchus kisutch). International Association for Genetics in Aquaculture, Genetics in Aquaculture VIII, 9-15 November, 2003: 102, Puerto Varas, Chile.

Cherfas, N.B. 1981. Gynogenesis in fishes. In: Genetic bases of fish selection, Editor Kirpichnikov, V.S. Springer-Verlag, Berlin, heidelberg, New York, 412 pp.

Chevassus, B. 1976. Variabilité et heritabilité des performances de croissance chez la truite arc-en-ciel *(Salmo gardneri Richardson).* Ann. Génét. Sél. Anim., 18: 273-283.

Chevassus, B. 1979. Hybridization in salmonids: Results and prespectives. Aquaculture, 17, 113-128.

Chevassus, B. and Dorson, M. 1990. Genetics of resistance to disease in fishes. Aquaculture, 85: 83-107.

Chourrout, D. 1980. Thermal induction of diploid gynogenesis and triploidy in the eggs of the rainbow trout *(Salmo gairdneri Richardson).* Reproduction Nutrition Development, 20 (3A): 727-733.

Chourrout, D. 1982a. Gynogenesis caused by ultraviolet irradiation of salmonid sperm. Journal of Experimental Zoology, 223: 175-181.

Chourrout, D. 1982b. Tetraploidy induced by heat shocks in the rainbow trout (*Salmo gairdneri* R.). Reprod. Nutr. Develop., 2: 569-574

Chourrout, D. 1984. Pressure-induced retention of second polar body and suppression of first cleavage in rainbow trout; production of all-triploids, all-tetraploids, and heterozygous and homozygous diploid gynogenetics. Aquaculture, 36: 111-126.

Clayton, G.A. and Robertson, A. 1957. An experimental check on quantitative genetical theory. II. The long-term effects of selection. J. genet., 55, 152-170.

Clifford, S.L., McGinnity, P. and Ferguson, A., 1998a. Genetic changes in an Atlantic salmon population resulting from escaped juvenile farm salmon. J. Fish Biol., 52: 118-127.

Clifford, S.L., McGinnity, P. and Ferguson, A., 1998b. Genetic changes in Atlantic salmon (*Salmo salar*) populations of northwest Irish rivers resulting from escapes of adult farm salmon. Can. J. Fish. Aquat. Sci., 55: 358-363.

Cnaani, A., Hallerman, E. M., Ron, M., Weller, J. L., Indelman, M., Kahi, Y., Gall, G. A. E. and Hulata, G. 2003. Detection of a chromosomal region with two quantitative trait loci, affecting cold tolerance and fish size, in an F_2 tilapia hybrid. Aquaculture, 223: 117-12 8.

Cochrane, W.G. 1951. Improvement by means of selection. Proc. 2nd Berkely Symp. Math. Stat. Prob. Univ. of California Press, Berkeley. Pp449-470.

Comstock, R.E., Robinson, H.F. and Harvey, P.H. 1949. A breeding procedure designed to make maximum use of both general and specific combining ability. J.Am. Soc. Agron. 41,, 360-367.

Connell, J. (1980) Control of fish quality. Fishing News books Ltd, Farnham, Surry, England, 222 pp. University of Agricultural Sciences, 1-13.

Cooper, E.L. 1961. Growth of wild and hatchery strains of brook trout. Trans. Am. Fish Soc., 90: 424-438.
Crawford, R.D. 1990. Poultry breeding and genetics. Elsevier, Amsterdam-Oxford-New York-Tokyo. 1123 pp.
Crenshaw, J.W.Jr., Heffernan, P.B. and Walker, R.L. 1991. Heritability for growth rate in the southern bay schallop. Argopecten irradians concentricus (Say 1822). J. Shellfish Res. 10(1): 55-63.
Crenshaw, J.W.Jr., Heffernan, P.B. and Walker, R.l. 1996. Effect of grouout density on heritability of growth rate in the northern quahog, Mercenaria mercenaria (Linnaeus,1758). Journal of Shellfish Research, 15: 341-344.
Crollius, H.R., Jaillon, O., Bernot, A., Dasilva, C., Bouneau, L., Fisher, C., Fizames, C., Wincker, P., Brottier, P., Quetier, F., Saurin, W. and Weissenbach, J. 2000. Estimate of human gene number provided by genomewide analysis using Tetradon nigroviridis DNA sequence. Nature America Inc. 25: 235-238.
Crow, J.F. and Kimura, M. 1970. An Introduction to population genetics theory, Harper and Row, New York, USA.
Crozier, W.W. 1993. Evidence of genetic interaction between escaped farmed salmon and wild Atlantic salmon (*Salmo salar*) in a Northern Irish river. Aquaculture 113: 19-29.
Darvasi, A. and Soller, M. 1992. Selective genotyping for determination of linkage between a marker locus and a quantitative trait locus. Theoretical and Applied Genetics, 85: 353-359.
Darvasi, A. and Soller, M. 1994. Selective DNA pooling for determination of linkage between a molecular marker and a quantitative trait locus. Genetics, 138: 1365-1373.
Darwin, C.R. 1859. The origin of species. The Modern Library, New York.
Davidson, W. S. and Koop, B. F. 2003. GRASP: Genomic research on Atlantic Salmon Project. Proc. Genetics in Aquacutlure, VIII. Page 17.
Davis, C.V. 2000. Estimation of narrow-sense heritability for larval and juvenile growth traits in selected and unselected sub-lines of eastern oysters, *Crassostrea virginica.* Journal of Shellfish Research, 19: 613.
Demster, E.R. and Lerner, I.M. 1950. Heritability of treshold characters. Genetics, 35: 212-236.
Devlin, R.H., Biagi, C.A., Yesaki, T.Y., Smailus, D.E. and Byatt, J.C. 2001. Growth of domesticated transgeneic fish. Nature, 409: 781-782.
Dey, M.M. and Eknath, A.E. 1996. Current trends in the Asian tilapia industry and the significance of genetically improved tilapia breeds. In: Sustainable Aquaculture: (Edited) Nambiar, K.P.P. and Singh, T. Proceedings of INDOFISH – Aquatech 1996, International Conference on Aquaculture, Kuala Lumpur, Malaysia, 25-27 Sept. 1996: 59-78.
Dickerson, G.E. 1952. Inbred lines for heterosis tests? In: Gown, J.W. (ed.), Heterosis, 330-351 pp., Iowa State College, Maes Iowa.
Diter, A., Quillet, E. and Chourrout, D. 1993. Suppression of first egg mitosis induced by heat shocks in the rainbow trout. J. Fish Biol., 42: 777-786
Donaldson, L.R. and Olson, P.R. 1955. Development of rainbow trout broodstock by selective breeding. Trans. Amer. Fish. Soc., 85: 93-101.
Doyle, R.W. 1983. An approach to the quantitative analysis of domestication selection in aquaculture Aquaculture, 33: 167-185.
Doyle, R.W. and Herbinger, C. 1994. The use of DNA fingerprinting for high-density, within-family selection in fish breeding. Proc. 5th World Congress on Genetics Applied to Livestock Production, Guelph, Canada, 19: 23-27.
Doyle, R.W. and Talbot, A.J. 1986a. Artificial selection on growth and correlated selection on competitive behavior in fish. Can. J. Fish. Aquat. Sci., 43: 1059-1064.
Doyle, R.W. and Talbot, A.J. 1986b. Effective population size and selection in variable aquaculture stocks. Aquaculture, 57: 27-35.
Dunham, R.A. 1986. Selection and crossbreeding responses for cultured fish. Proc. 3rd World Congress on Genetic Applied to Livestock Production, X: 391-400.
Dunham, R.A. 1987. American catfish breeding programs. Proc. World Symp. on selection, hybridisation, and genetic enginering in aquaculture, Bordeaux 27-30 May, 1986, Berlin, 11: 407-416.
Dunham, R.A., Brummett, R.E., Ella, M.O. and Smitherman, R.O. 1990. Genotype-environment interactions for growth of blue, channel and hybrid catfish in ponds and cages at varying densities. Aquaculture, 85: 143-151.
Dunham, R.A., Warr, G.W., Nichols, A., Duncan, P.L., Argue, B., Middleton, D., and Kucuktas, H. 2002. Enhanced bacterial disease resistance of transgenic Channel catfish *Ictalurus punctatus* possessing cecropin genes. Mar. Biotechnol., 4: 338-344.

Edwards, D. and Gjedrem, T. 1979. Genetic variation in survival of brown trout eggs, fry and fingerlings in acidic water. SNSF - project, Norway, FR 16/79, 28 pp.

Ehlinger, N.F. 1977. Selective breeding of trout for resistance to furunculosis. New York Fish and Game Journal, 24: 25-36.

Eide, D.M., Linder, R.D., Fjalestad, K., Larsen, H.J.S., Røed, K.H. and Strømsheim, A. 1994. Genetic variation in antibody response to diphtheria toxoid in Atlantic salmon and rainbow trout. Aquaculture, 127: 103-113.

Eisen, E.G. 1980. Conclutions from long-term selection experiments with mice. Z. Tierz. Zuchtgsbiol., 97: 305-319.

Eknath, A.E. and Acosta, B.O. 1998. Genetic Improvement of Farmed Tilapias (GIFT) Project: final report, March 1988 to December 1997. ICLARM. Makati City.

Eknath, A.E., Bentsen, H.B., Gjerde, B. Tayamen, M.M., Abella, T.A., Gjedrem, T. and Pullin, R.S.V. 1991. Approaches to national fish breeding programs: Pointers from tilapia pilot study. NAGA, The ICLARM Quarterly No 723:10-12.

Eknath, A.E., Dey, M.M., Rye, M., Gjerde, B., Abella,T.A.,Sevilleja, R., Tayamen, M.M., Reyes, R.A. and Bentsen, H.B. 1998. Selective breeding of nile tilapia for Asia. Proceedings of the 6th World Congress on Genetics Applied to Livestock Production, 27: 89-96.

Eknath, A.E., Tayamen, M.M., Palada-de Vera, M.S., Danting, J.C., Reyes, R.A., Dinosio, E.E., Capili, J.B., Bolivar, H.L., Abella, T,A., Circa, A.V., Bentsen, H.B., Gjerde, B., Gjedrem, T. and Pullin, R.S.V. 1993. Genetic improvement of farmed tilapia: the growth performance of eight strains of *Oreochromis niloticus* tested in different farm environments. Aquaculture, 111: 171-188.

El-Ibiary, H. M. and Joyce, J. A. 1978. Heritability of body traits, dressing weight and lipid content in channel catfish. J. Anim. Sci., 47: 82-88.

Ellegren, H. and Fridolfsson, A.K. 2003. Sex-Specific mutation rates in salmonid fish. J. Mol. Evol., 56: 458-463.

Elvingson, P. 1992. Estimates of phenotypic and genetic parameters of carcass lipid content on rainbow trout *(Oncorhynchus mykiss)* based on specific gravity measurements. Dr. Thesis. Swedish University of Agricultural Sciences, 1-10.

Elvingson, P. and Nilsson, J. 1992. Phenotypic and genetic parameters of body and compositional traits in Arctic char *(Salvelinius alpinus).* Dr. Thesis, Swedish University of Agricultural Sciences, 1-13.

Embody, G.C. and Hyford, C.D. 1925. The advantage of rearing brook trout fingerlings from selected breeders. Trans. Amer. Fish. Soc., 55: 135-138.

Enfield F.D. 1974. Recurrent selection and response plateaus. In: Proceedings of the First World Congress on Genetics Applied to Livestock Production, Madrid, 1:365-371.

Enfield, F.D. 1979. Long term effects of selection; the limits to response. In proceedings of a symposium on "Selection experiments in laboratory and domestic animals". Commonwelth Agric. Bureaux, pp. 69-86.

Enfield F.D. 1988. New sources of variation. In: Proceedings of the Second International Conference on Quantitative Genetics, pp. 215-218. Sinauer, Sunderland, MA.

Estoup, A., Gharbi, K., SanCristobal, M., Chevalet, C., Haffray, P. and Guyomard, R. 1998. Parentage assignment using microsatellites in turbot *(Scophthalmus maximus*) and rainbow trout *(Oncoryhnchus mykiss)* hatchery populations. Can. J. Fish Aquat. Sci., 55: 715-725.

Evans, F., Matson, S., Brake, J. and Langdon, C. 2004. The effect of inbreeding on performance traits of adult Pacific oysters *(Crassostrea gigas).* Aquaculture, 230: 89-98.

Falconer, D.S. 1955. Patterns of response in selection experiments with mice. Cold Spring Harbor Symp. Quant. Biol., 20: 178-196.

Falconer, D.S. 1960. Introduction to Quantitative Genetics. Oliver and Boyd, Edinburgh, 365 pp.

Falconer, D.S. 1965. The inheritance of liability to certain diseases, estimated from the incidence among relatives. Ann. Hum. Genet., 29: 51-76.

Falconer, D.S., 1981. Introduction to quantitative genetics. Second edition. 340 pp., Longman, London and New York.

Falconer, D.S. and Mackay, T.F.C. 1996. Introduction to quantitative genetics. Longman, Essex CM20 2JE, England, 464 pp.

FAO, 2003. FAO Yearbook. Fishery statistics, Aquaculture production. Vol. 92/2 2001, 86pp.

Feist, G., Yeoh, C.G., Fitzpatrick, M.S. and Schreck, C.B. 1995. The production of functional sex-reversed mals rainbow trout with 17α-methyltestosterone and 11β-hydroxyandrostenedione. Aquaculture, 131: 145-152.

Ferguson, A., Taggart, J.B., Prodöhl, P.A., McMeel, O., Thomson, C., Stone, C., McGinnity, P. and Hynes, R.A. 1995. The application of molecular markers to the study and conservation of fish populations, with special reference to Salmo. J. Fish Biol., 47 (Suppl. A): 103-126.

Fernando, R. and Grossman, M. 1988. Marker assisted selection using best linear unbiased prediction. Genetics Selection Evolution, 21: 467-477.

Fevolden, S.E., K.H. Røed, and B. Gjerde. 1994. Genetic Components of Post-Stress Cortisol and Lysozyme Activity in Atlantic Salmon: Correlations to Disease Resistance. Fish & Shellfish Immunology, 4:507-519.

Fimland, E. 1979. The effect of selection on additive genetic parameters. Zeitschrift fur Tierzuchtung und Zuchtungsbiologie, 96: 120-134.

Fjalestad, K.T., Gjedrem, T., Carr, W.H. and Sweeney, J.N. 1997. Final report: The shrimp breeding program Selective breeding of *Penaeus vannamei*. AKVAFORSK, Report no. 17/97, 85 pp.

Fjalestad, K.T., Gjedrem, T. and Gjerde, B. 1993. Genetic improvement of disease resistance in fish: an overview. Aquaculture, 111: 65-74.

Fjalestad, K.T., Gjedrem, T, Lund, T. Røed, K.H. and Refstie, T. 1994. Immune parameters as indirect selection criteria against furunculosis. In manuscript.

Fjalestad, K.T., Larsen, H.J.S., and Røed, K.H. 1996. Antibody response in Atlantic salmon *(Salmo salar)* against Vibrio anguillarum and Vibrio salmonicida O-antigens: Heritabilities, genetic correlations and correlations with survival. Aquaculture, 145:77-89.

Fleming, I.A., Jonsson, B., Gross, M.R. and Lamberg, A. 1996. An experimental study of the reproductive behaviour and success of farmed and wild Atlantic salmon (Salmo salar). J. Appl. Ecology, 33, 893-905.

Fleming, I.A., Lamberg, A., and Jonsson, B. 1997. Effects of early experience on the reproductive performance of Atlantic salmon. Behavioral Ecology, 8, 470-480.

Flynn, F.M.O., Bailey, J.K. and Friars, G.W. 1999. Response to two generations of index selection in Atlantic salmon *(Salmo salar).* Aquaculture, 173: 143-147.

Fosil. L. and Chourrout, D. 1992. Chromosome doubling by pressure treatments for tetraploidy and mitotic gynogenesis in rainbow trout, Oncorhynchus mykiss (Walbaum): RE-examination and improvements. Aquacult. Fish. Manage., 23: 567-575.

Frankel, O.H. and Soulé, 1981. Conservation and evaluation. Cambridge Univ. Press, Cambridge.

Frankenhauser, G. 1945. The effect of changes in chromosome numbers on amphibian development. Q. Rev. Biol., 20: 20-78.

Fredrich, F. and Reuter, W. 1982. Hinweise zur weiteren Verbesserung der kunstichen Vermehrung der Regenbogenforelle Salmo gairdnerii R. Binnenfisch. DDR 29 (5): 146-150.

Freeman, A.M. 1993. The measurement of environmental and resourde values: Theory and methods. Resources for the future. Washington DC, 516 pp.

Friars, G.W. 1993. Breeding Atlantic salmon: A primer. ASF, St. Andrews, NB, 13 pp.

Friars, G. W., Bailey, J.K. and Saunders, R.L. 1979. Considerations of a method of analysing diallel crosses of Atlantic salmon. Can. J. Genet. Cytol, .21: 121-128.

Galbreath, P.F. and Thorgaard, G.H. 1995. Saltwater performance of all female triploid Atlantic salmon. Aquaculture, 138: 77-85.

Gall, G.A.E. 1972. Phenotypic and genetic components of body size and spawning performance. In Progress in Fishery and Food Science, R.W. Moor (Editor). Univ. of Washington Publications in fisheries, New Series Vol: 5 Univ. of Washington, Seattle, Wa.

Gall, G.A.E. 1974. Influence of size of eggs and age of female on hatchability and growth in rainbow trout. Calif. Fish. Game, 60: 26-35.

Gall, G.A.E. 1975. Genetic of reproduction in domesticated rainbow trout. J. Anim.Sci., 40, 19-28.

Gall, G.A.E. 1990. Basis for evaluating breeding plans. Aquaculture, 85: 125-142.

Gall, G.A.E. and Bakar, Y. 2002. Application of mixed-model techniques to fish breed improvement:analysis of breeding-value selection to increase 98-day body weight in tilapia. Aquaculture, 212: 93-113.

Gall, G.A.E., Baker, Y. and Famula, T. 1993. Estimating genetic change from selection. Aquaculture, 111: 75-88.

Gall, G.A.E. and Huang, N. 1988a. Heritability and selection schemes for rainbow trout: body weight. Aquaculture, 73: 43-56.

Gall, G.A.E. and Huang, N. 1988b. Heritability and selection schemes for rainbow trout: Female reproductive performance. Aquaculture, 73: 57-66.

Gardner, E.J. and Snustad, D.P. 1981. Principles of genetics. John Wiley and Sons, New York, Chichester, Brisbane, Toronto, Singaphore, 611 pp.

Garcia, D., Carleos, C., Parra, D. and Canon, J. 2002. Sib-parentage testing using molecular markers when parents are unkonwn. Animal Genetics, 33: 364-371.
Gervai, J., Peter, S., Nagy, A., Horvath, L. and Csanyi, V. 1980. Induced triploidy in carp, *Cyprinus carpio L*. J.Fish. Biol., 17: 667-671.
Gjedrem, T. 1967. Selection indexes compared with single trait selection. 1. The efficiency of including correlated traits. Acta Agric. Scand., 17: 263-268.
Gjedrem, T. 1976. Genetic variation in tolerance of brown trout to acid water. SNSF - project, Norway, FR 5/76, 11 pp.
Gjedrem, T. 1979a. Selection for growth rate and domestication in Atlantic salmon. Z. Tierzuchtg. Zuchtgsbiol., 96: 56-59.
Gjedrem, T. 1979b. Farming of salmon and trout. (Oppdrett av laks og aure). Landbruksforlaget, Oslo, 332 pp.
Gjedrem, T. 1983a. Genetic variation in quantitative traits and selective breeding in fish and shellfish. Aquaculture, 33: 51-72.
Gjedrem, T. 1983b. Possibilities for genetic changes in the salmonides. Roczniki nauk rolniczych, 3: 65-78.
Gjedrem, T. 1985. Improvement of productivity through breeding schemes. GeoJournal, 10(3): 233-241.
Gjedrem, T. 1992. Breeding plan for rainbow trout. Aquaculture, 100: 73-83.
Gjedrem, T. 1995a. Genetics and breeding for aquaculture. (Genetikk og avlslære for akvakultur). Landbruksforlaget, 112 pp.
Gjedrem, T. 1995b. Diseases and parasites in fish. Genetic improvement of resistance. GIFT, Final Report (March 1988 to December 1997). ICLARM publication: Attachment 11: 27pp.
Gjedrem, T. 1997a. Flesh quality improvement in fish through breeding. Aquaculture International, 5: 197-206.
Gjedrem, T. 1997b. Selective breeding to improve aquaculture production. World Aquaculture, 28(1): 33-45.
Gjedrem, T. 1998. Developments in fish breeding and genetics. Acta Agric. Scand., Sect. A. Animal Sci., 28: 19-26.
Gjedrem, T. and Aulstad, D. 1974. Selection experiments with salmon. I. Differences in resistance to vibrio disease of salmon parr (Salmo salar). Aquaculture, 3: 51-59.
Gjedrem, T. and Fimland, E. 1995. Potential benefits from high health and genetically improved shrimp stocks. Proc. of the Special Session on Shrimp Farming: Swimmingthrough troubled water. The World Aquaculture Society, February 1-4 1995, San Diego, California: 60- 65.
Gjedrem, T. and Fjalestad, K.T. 1997. Body weight distribution of Atlantic salmon parr during domestication. AKVAFORSK, manuscript, 14pp.
Gjedrem, T. and Gjøen, H.M. 1995. Genetic variation in susceptibility of Atlantic salmon, *Salmo salar L*., to furunculosis, BKD and coldwater vibriosis. Aquaculture Research, 26: 129-134.
Gjedrem, T., Gjøen, H.M. and Gjerde, B., 1991. Genetic origin of Norwegian farmed Atlantic salmon. Aquaculture, 98, 41-50.
Gjedrem. T., Gjøen, H.M., Hardjamulia, A., Sudarto, Ir., Widiaty, A., Gustiano, R., Kristanto, A.H., Emmawati, L. and Hadie, W. 1997. Breeding plan for Nile tilapia in Indonesia: Individual (mass) selection. INGA, ICLARM, Report No. 4. 11 pp.
Gjedrem, T, Refstie, T. and Gjerde, B. 1987. A review of quantitative genetic research in salmonids at AKVAFORSK. Second Int. Conference on Quantitative Genetics: 527-535.
Gjedrem, T., Salte, R. and Gjøen, H.M. 1991. Genetic variation in susceptibility of Atlantic salmon to furunculosis. Aquaculture, 97:1-6.
Gjerde, B. 1984a. Response to individual selection for age at sexual maturity in Atlantic salmon. Aquaculture, 38: 229-240.
Gjerde, B. 1984b. Variation in semen production of farmed Atlantic salmon and rainbow trout. Aquaculture, 40: 109-114.
Gjerde, B. 1986. Growth and reproduction in fish and shellfish. Aquaculture, 57:37-55. Gjerde, B. 1987. Predicting carcass composition of rainbow trout by computerised tomography. J. Anim. Breed. Genet., 104: 121-136.
Gjerde, B. 1987. Predicting carcass composition of rainblw trout by computerized tpmography. J. Anim. Breed. Genet., 104:121-136.
Gjerde, B. 1988. Complete diallele cross between six inbred groups of rainbow trout, *Salmo gairdneri.* Aquaculture, 75: 71-87.
Gjerde, B. 1991. Breeding theory, salmon and rainbow trout. (Avlslære, laks og regnboge aure). Landbruksbokhandelen, Norway. ISBN 82-557-0374-8.

Gjerde, B. 2003. Genetic markers for practical assignement in fish breeding programs. International Association for Genetics in Aquaculture, Genetics in Aquaculture VIII. 9-15 November 2003: 75, Puerto Varas, Chile

Gjerde, B and Gjedrem, T. 1984. Estimates of phenotypic and genetic parameters for carcass traits in Atlantic salmon and rainbow trout. Aquaculture, 36: 97-110.

Gjerde, B., Gjøen, H.M. and Villanueva, B. 1996. Optimum designs for fish breeding programmes with constrained inbreeding. Mass selection for normally distributed traits. Livestock Production Science, 47: 59-72.

Gjerde, B., Gunnes, K. and Gjedrem, T. 1983. Effect of inbreeding on survival and growth rate. Aquaculture, 34: 327-332.

Gjerde, B. and Korsvoll, A. 1999. Realized selection differentials for growth rate and early sexual maturity in Atlantic salmon. Abstracts, Aquaculture Europe 99, Trondheim, Norway, August 7-10: 73-74.

Gjerde, B., Mahapatra, K.D., Reddy, P.V.G.K., Saha, J.N., Jana, R.K. and Rye, M. 2004. Genetic parameters for growth of rohu carp (Labo rohita). Aquaculture, in manuscript.

Gjerde, B. and Reddy, P.V.G.K. 1996. Results from the project selective breeding of rohu. Final Workshop on Selective Breeding of Rohu. An Indo-Norwegian Collaborative Project CIFA and AKVAFORSK, 21st and 22nd March, 1996, CIFA, Orissa, India.

Gjerde, B., Reddy, P.V.G.K., Rye, M., Mahapatra , K.D., Saha, J.N., Jana, R.K., Gupta, S.D., Sahu, M., Lenka, S. and P. Govindassamy, P. 1999. Average heterosis for growth and survival in two diallel crosses with four stocks of rohu carp, (Labeo rohita). European Aquaculture Society, Special.

Gjerde, B. and Refstie, T. 1984. Complete diallel cross between five strains of Atlantic salmon. Livest. Prod. Sci., 11: 207-226.

Gjerde, B. and Rye, M. 1997. Design of breeding programs in aquaculture species - Possibilities and constraints. Genetics and breeding of Mediterranean aquaculture species. Proc. of the seminar of the CIHEAM Network on Technology of Aquaculture in the Mediterranean (TECAM), jointly organized by CIHEAM and FAO, Zaragoza, Spain. Options Mediterr. 34:181-192. ISSN:1022-1379.

Gjerde, B., Rører, J.E., Lein, I., Stoss, J. and Refstie, T. 1997. Heritability for body weight in farmed turbot. Aquaculture International, 5: 175-178.

Gjerde, B. and Schaeffer, L.R. 1989. Body traits in rainbow trout. II. Estimates of heritabilities and phenotypic and genetic correlations. Aquaculture, 80: 25-44.

Gjerde, B., Terjesen, B.F., Barr, Y., Lein, I. and Thorland, I. 2004. Genetic variation for juvenile growth and survival in Atlantic cod (Gadus morhua). Aquaculture, 236:167-177.

Gjerde, B., Villanueva, B. and Bentsen, H.B., 2002. Opportunities and challenges in designing sustainable fish breeding programs. In: Proc. of the 7th World Congress on Genetics Applied to Livestock Production, Montpellier, France, 30: 461-468.

Gjøen, H.M. 1996. Studies on estimation of genetic parameters and design of optimal breeding schemes for survival traits in Atlantic salmon. Ph thesis at Agricultural University of Norway. 79 pp.

Gjøen, H. M. and Bentsen, H.B. 1997. Past, present and future of genetic improvement in salmon aquaculture. ICES J. Marine Sci., 54: 1009-1014.

Gjøen, H.M., Gjedrem, T., Hardjamulia, A., Sudarto, Ir., Widiaty, A., Gustiano, R., Kristanto, A.H., Emmawati, L. and Suparta;M. 1997. Breeding plan for common carp in Indonesia: Combined multitrait selection. INGA, ICLARM, Report No.5, 12 pp.

Gjøen, H.M. and B. Gjerde. 1998. Comparing breeding schemes using individual phenotypic values and BLUP breeding values as selection criteria. Proc.6th World Congress on Genetics Applied to Livestock Production 111-114.

Gjøen, H.M., Storebakken, T., Austreng, E. and Refstie, T. 1993. Genotypes and nutrient utilization. Fish Nutrition in Practise, Blarritz (France), June 24-27, 1991. Ed. INRA, Paris (Les Colloques, n°61), 19-26 pp.

Goddard, M. E. and Hayes, B. J. 2002. Optimisation of selection response using molecular data. Proc. 7th World. Congr. Genet. Appl. Livest. Prod. Montpellier, France.

Godin, D.M., Carr, W.H., Hagino, G., Segura, F., Sweeney, J.N. and Blankenship, L. 1996. Evaluation of a fluorescent elastomer internal tag in juvenile and adult shrimp *Penaeus vannamei.* Aquaculture, 139: 243-248.

Gold, J.R. 1979. Cytogenetics. In "Fish Physiology" (Editor: Hoar, W.S. Randall, D.J. and Brett, J.R.), 8: 353-405. Academic Press, New York.

Gold, J.R. and Avise, J.C. 1976. Spontaneous triploidy in the California roach *Hesperoleucus symmetricus* (Pisces: *Cyprinidae*). Cytogenet. and Cell Genet., 17: 144-149.

Golovinskaya, K.A. 1968. Genetics and selection of fish and artificial gynogenesis of the carp *(Cyprinus carpio*). FAO Fisheries Reports, (44)4: 215-222.

Gow, R.S. and Fairfull, R.W. 1990. Genetics controls in selection. In "Poultry Breeding and Genetics", pp. 935-954, Editor R. D. Crawford, Elsevier, Amsterdam, Netherlands.

Grundy, B., Villanueva, B. and Woolliams, J.A. 1998. Dynamic selection procedures for constrained inbreeding and their consequences for pedigree development. Genet. Res., Camb., 72:159-168.

Grundy, B., Villanueva, B. and Woolliams, J.A. 2000**.** Breeding and genetics - Dynamic selection for maximizing response with constrained inbreeding in schemes with overlapping generations. Anim. Sci., 70:373-382.

Gunnes, K. and Gjedrem, T. 1978. Selection experiments with salmon. IV. Growth of Atlantic salmon during two years in the sea. Aquaculture, 15: 19-23.

Gunnes, K. and Gjedrem, T. 1981. A genetic analysis of body weight and length in rainbow trout reared in sea water for 18 months. Aquaculture, 24: 161-174.

Gunnes, K. and Refstie, T. 1980. Cold branding and fin-clipping for marking of salmonids. Aquaculture, 19: 295-299.

Gunsett, F.C. 1986. Problems associated with selection for traits defined as a ratio of two component traits. in: Proc. of the 3rd World Congr. Genetics Applied to Livestock Prod., Univ. Nebraska, Lincoln, Nebraska, USA., 11: 437-442.

Guo, X., DeBrosse, G.A. and Allen, S.K. 1996. All-triploid Pacific oysters (Crassostrea gigas Thunberg) produced by mating tetraploids and diploids. Aquaculture, 142: 149-164.

Gupta, M.V. and Acosta, B.O. 2001. Networking in aquaculture genetics research, p.1-5. In. M.V. Gupta and B.O. Acosta (eds.), Fish genetics research in member countries and institutions at the International Network on Genetics in Aquaculture. ICLARM Conf. Proc. Penang.

Hadley, N.H., Dillon, R.F.Jr. and Manzi, J.J. 1991. Realized heritability of growth rate in the hard clam Mrcenaris mercenaria. Aquaculture, 93: 109-119.

Haldane, J.B.S. 1919. The combination of linkage values and the calculation of distances between linked factors. J. Genet., 8: 299–309.

Haldane, J.B.S. 1946. The interaction of nature and nurture. Ann. Eugen. 13: 197.

Hale, E.B. 1969. Domestication and evolution of behavior. In: E.S.E. Hafez (Editor), The behavior of domestic animals. Ballier, Tindall and Castell, London, 22-42 pp.

Haley, L.E. and Newkirk, G.E. 1982. The genetics in growth rate of Crassostrea virginiia and Ostrea edulis. Malaclogia, 22: 399-401.

Haley, L.E., Newkirk, G.F., Waugh, D.W. and Doyle, R.W. 1975. A report on the quantitative genetics of growth and survivalship of the American oyster, *Crassostera virginica* under laboratory conditions. In: 10th Eur. Symp. Mar. Biol. Ostend, Belium, 17-23 September 1975, 1: 221-228.

Hallermann, E.M. and Beckmann, J.S. 1988. DNA-level polymorphism as a tool in fisheries science. Can. J. Fish. Aquat. Sci., 45: 1075-1087.

Hammond, J. 1947. Animal breeding in relation to environmental conditions. Biol. Rev., 22: 195.

Hammond, K. 1994. Conservation of domestic diversity: global overview. In: Proceedings of the 5th World Congress on Genetics Applied to Animal Production, Guelph, Canada, 21: 423-430.

Hanke, A.R., Friars, G.W., Saunders, R.L. and Terhune, J.M. 1989. Family x photoperiod interaction on growth in juvenile Atlantic salmon, Salmo salar. Genome, 32(6): 1105-1112.

Hanocq, E., Boichard, D. and Foulley, JL. 1996. A simulation study of the effect of connectedness on genetic trend. Genet Sel Evol., 28: 67-82.

Hard, J.J. and Herschberger, W.K. 2002. Research on quantitative genetic consequences of captive broodstock programs for Pacific salmon populations. Annual report to Bonneville Power Administration, P.O. Box 3621, Portland, OR 97208-3621.

Hardy, G.H. 1908. Discussion and correspondence: Mendelian proportions in a mixed population. Science, 28: 49-50.

Harvey, W.R. 1964. Computing procedures for a generalised least square analysis program. Paper presented at the Analysis of Variance Conference, Colorado State University, Fort Collins, Colo., 50 pp. (mimeographed).

Haskin, H.H. and Ford, S.E. 1979. Development of resistance to *Minchinia nelsoni* (MSX) mortality in laboratory rearred and native oyster stocks in Deleware Bay. Mar. Fish. Rev., 41: 54-63.

Hayes, B. and Goddard, M.E. 2001. The distribution of the effect of genes affecting quantitative traits in livestock. Genet. Sel. Evol., 33: 209-229.

Hazel, L.N. 1943. The genetic basis for constructing selection indexes. Genetics, 28: 476-490.

Hazel, L. N. and Lush, J. L. 1942. The efficiency of three methods of selection. J. Hered., 33: 393-399.
He, M., Lin, Y., Shen, Q.,Jianxin,J. and Jiang, W. 2000. Production of tetraploid pearl oyster (Pinctada martensii Dunker) by inhibiting the first polar body in eggs from triploids. J.Shellfish Res., 19: 147-151
Hadley, N.H., Dillon, R.T.Jr. and Manzi, J.J. 1991. Realized heritability of growth rate in the hard clam Mercenaris mercenaria. Aquaculture, 93: 109-119.
Hackett, P.B. 1993. The molecular biology of trasgenic fish. In: hochachka and Mommsen (eds.) Bimchemistry and molecular biology of fishes. Vol. 2, Elsevier Sci. Publ.
Heffernan, P.B., Walker,R.L. and Ryan, M. 1993. Second heritability estimate of growth rate in the southern bay scallop. Argopecten irradians concentricus (Say 18229. J. Shellfish Res., 12(1): 151.
Heggberget; T.G., Johnsen, B.O., Hindar, K., Jonsson, B., Hansen, L.H., Hvidsten, N.A. and Jensen, A.J. 1993. Interactions between wild and cultured Atlantic salmon: a review of the Norwegian experience. Fish. Res., 18: 123-146.
Helland, S.J., Grisdale-Helland, B. and Nerland, S. 1996. A simple method for measurement of daily feed intake of groups of fish in tanks. Aquaculture, 139: 157-163.
Henderson, C.R. 1949. Estimation of changes in herd environment (abstract). Journal of Dairy Science, 32: 709.
Henderson, C.R. 1950. Estimation of genetic parameters. Annals of Mathematical Statistics, 21:309-.
Henderson, C.R. 1953. Estimation of variance and covariance components. Biometrics, 9: 226-252.
Henderson, C.R. 1963. Selection index and expected genetic advance. NAS-NRC. 982.
Henderson, C.R. 1973. Sire evaluation and genetic trends. In: Proceedings of the Animal Breeding and Genetics Symposium in Honour of J.L. Lush. American Society for Animal Science, Blackburgh, Champaign, Illinois, pp. 10-41.
Henderson, C.R. 1975. Best linear unbiased estimation and prediction under a selaction model. Biometrics, 31: 423-437.
Henderson, C.R. 1976. A simple method for computing the inverse of a numerator relationship matrix used in predicting of breeding values. Biometrics, 32: 69-83.
Henderson, C.R. 1984. Application of Linear Models in Animal Breeding. University of Guelph Press, Guelph, Canada.
Heringstad, B. 2001. Liability to clinical mastitis. J. Dairy Sci., 84 (11): 2337-2346.
Hershberger, W.k., Meyers, J.M., McAuley, W.C. and Saxton, A.M. 1990. Genetic changes in growth of coho salmon *(Oncorhynchus kisutch)* in marine netpens, produced by ten years of selection. Aquaculture, 85: 187-197.
Hertwig, O. 1911. Die radiumkrankheit tierischer Kiemzellen. Archiv fur Mikroscopische Anatomie und Entwicklungsmechanik, 77: 1-97.
Hetzel, D.J.,S., Crocos, P.J., Davis, G.P., Moore, S.S. and Preston, N.C. 2000. Response to selection and heritability for growth in the Kuruma prawn, Penaeus japanicus. Aquaculture, 181: 215-223.
Hew, C.L., Fletcher, G.L. and Davies, P.L. 1995. Transgenic salmon: tailoring the genome for food production, 1995. J. Fish. Biol., 47 (Suppl. A), 1-19.
Hickling, C.F. 1962. Fish culture. Faber and Faber, London.
Hindar, K., Ryman, N. and Utter, F. 1991. Genetic effects of cultured fish on natural fish populations. Canadian J. Fish. and Aquat. Sci., 48: 945-957.
Holmefjord, I. 1986. DNA-measuring and induction of triploidy in Atlantic salmon and rainbow trout. (DNA-måling og induksjon av triploidi hos laks og regnbueørret). 1. Flow cytometric DNA-analysis of some salmonid species, hybrids and induced triploids. 2. Optimization of heat shock treatment for induction of triploidy in Atlantic salmon (Salmo salar) compared to rainbow trout *(Salmo gairdneri)*. Msc. thesis, Department of Fisheries Biology, University of Bergen, Norway, 66 pp.
Howe, A. and Kocher, T. D. 2003. Comparitive mapping of a QTL for red body colour in Tilapia. Aquaculture Workshop, January 2003, San Diego.
Huang, C.M. and Liao, I.C. 1990. Response to mass selection for growth rate in Oreochromis niloticus. Aquaculture, 85: 199-205.
Hung, S.S.O., Storebakken, T., Cui, Y., Tian, L. and Einen, O. 1997. High-energy diets for white sturgeon, *Acipenser transmontanus* Richardson. Aquacult. Nutr., 3: 281-286.
Hubbs, C.L. and Hubbs, L.C. 1932. Apparent parthenogenesis in nature, in a form of fish of hybrid origin. Science, NY, 76: 628-630.
Huitfeldt-Kaas, H. 1946. Tribes of salmon in Norway. Nytt Mag. Naturvidensk. B 85: 115-159.
Hulata, G. 1995. A review of genetic improvement of the common carp *(Cyprinus carpio L.)* and other cyprinids by crossbreeding, hybridisation and selection. Aquaculture, 129: 143-155.

Hulata, G. 2002. Genetic manipulations in aquaculture: a review of stock improvement by classical and modern technologies Genetica, 111:155-173.

Hulata, G., Wohlfarth, G.W. and Halevy, A. 1986. Mass selection for growth rate in the Nile tilapia *(Oreochromis niloticus).* Aquaculture, 57: 177-184.

Hulata, G., Wohlfarth, G. and Rothbard, S. 1983. Progeny-testing selection of tilapia broodstocks producing all-male hybrid progenies- Preliminary results. Aquaculture, 33: 263-268.

Hunter, G.A. and Donaldson, E.M. 1983. Hormonal sex control and its application to fish culture, in Fish Physiology (eds Hoar, W.S., Randall, D.J. and Donaldson), Academic Press, New York, 9B: 223-303.

Hussain, M.G., Chatterji, A., McAndrew, B.J. and Johnstone, R. 1991. Triploidy induction in Nile tilapia, *Oreochromis niloticus L*. using pressure, heat and cold schocks. Theor. Appl. Genet., 81: 6-12.

Husain, M.G., McAndrew, B.J., Penman, D.J. and Sodsuk, P. 1994. Estimate gene-centromere recombination frequencies in gynogenetic diploids of *Oreochromis niloticus (L.)* using allozymes, skin colour and a putative sex-determination locus (SLD-2). In: genetics and evolution of aquatic organisms. Edited by A.R. Beaumont. Chapman and hall, London, UK. 502-508 pp.

Hussain, M.G., Rao, G..P.S., Humayun, N.M., Randall, C.F., Penman, D.J., Kime, D., Bromage, N.R. and McAndrew, B.J. 1995. Comparative performance of growth, biochemical composition and endocrine profiles in diploid and triploid *Oreochromis niloticus L.* Aquaculture, 138: 87-98.

Hyperdictionary, 2003. http://www.hyperdictionary.com/

Ibarra, A.M., Ramirez, J.L., Ruiz, C.A.,Cruz, P. and Avila, S. 1999. Realized heritabilities and genetic correlation after dual selection for total weight and shell width in catarina scallop (Argopecten ventricosus). Aquaculture, 175: 227-241.

Isaksson, A. 1988. Salmon ranching: a World Review. Aquaculture, 75: 1-33.

Iwamoto, R. N., Myers, J.M. and Hershberger, W. K. 1990. Heritability and genetic correlations for flesh coloration in pen-reared coho salmon. Aquaculture, 86: 181-190.

Jarayabhand, P. and Thavornyutikarn, M. 1995. Realized heritability estimation on growth rate of oyster, Succostrea cucullata born, 1778. Aquaculture, 138: 111-118.

Jarimopas, J. 1988. Realized response of Thai red tilapia to weight-specific selection for growth (3^{rd}-5^{th} generations). NACA, Bangkok, 13 pp.

Jarimopas, P. 1990. Realized response of Thai red tilapia to 5 generations of size-specific selection for growth. Proceedings of the second Asian fisheries forum, Tokyo, Japan, 17-22/4 1990, 519-522.

Jobling, M., Arnesen, A.M., Baadsvik, B.M., Christiansen, J.S. and Jørgensen, E.H. 1995. Monitoring feeding behaviour and food intake: methods and applications. Review article. Aquacult. Nutr., 1: 131-143.

Johnstone, R. 1985. Induction of triploidy in Atlantic salmon by heat shock. Aquaculture, 49: 133-139.

Johnston, W.L., Atkinson, J.L. and Glanville, N.T. 1994. A technique using sequential feedings of different coloured foods to determine food intake by individual rainbow trout, *Oncorhynchus mykiss*: effect of feeding level. Aquaculture, 120, 123-133.

Johnstone, R. and Youngson, A.F. 1984. The progeny of sex-inverted female Atlantic salmon *(Salmo salarL.).* Aquaculture, 37: 179-182.

Jonasson, J. 1984. Comparison of sexual maturation between diploid and triploid rainbow trout *(Salmo gairdneri).* Msc. thesis, Agricultural University of Norway, 43pp.

Jonasson, J. 1993. Selection experiments in salmon ranching: I. Genetic and environmental sources of variation in survival and growth in freshwater. Aquaculture, 109: 225-236.

Jonasson, J. 1994. Selection experiments in Atlantic salmon ranching. Phd thesis at Agricultural University of Norway, 97 pp.

Jonasson, J., Gjerde, B. and Gjedrem, T. 1997. Genetic parameters for return rate and body weight of sea-ranched Atlantic salmon. Aquaculture, 154:219-231.

Jonasson, J., Stefansson, S.E., Gudnason, A. and Steinarsson, A. 1999. Genetic variation for survival and shell length of cultured red abalone (Haliotis rufescens) in Iceland. J. of Shellfish Res., 18(2): 621-625.

Jonsson, N., Hansen, L.P. and Jonsson, B. 1991. Variation in age, size and repeat spawning of adult Atlantic salmon in relation to river discharge. J. Anim. Ecol., 60: 937-947.

Kackar, R.N. and Harville, D.A. 1981. Unbiasedness of two-stage estimation and prediction procedures for mixed linear models. Comm. in Stat. A10: 1249.

Kallman, K.D. 1962. Gynogenesis in the teleost, *Mollienesia formosa*, (Girard), with a discussion of the detection of parthegonesis in vertebrates by tissue transplantation. J. Genet., 58: 7-24.

Kallman, K.D. 1965. Genetics and geography of sex determination in the poeciliid fish. *Xiphophorus maculatus*. Zoologica, 50: 151-190.

Kanis, E., Reftie, T. and Gjedrem, T. 1976. A genetic analysis of egg, alevin and fry mortality in salmon *(Salmo salar),* sea trout *(Salmo trutta)* and rainbow trout *(Salmo gairdneri).* Aquaculture, 8:259-268.

Kause, A., Ritola, O., Paananen, T., Mäntysaari, E. and Eskelinen, U. 2002. Coupling body weight and its composition: a quantitative genetic analysis in rainbow trout. Aquaculture, 211:65-79.

Kause, A., Ritola, O., Paananen, T., Mäntysaari, E. and Eskelinen, U. 2003. Selection against early maturity in large rainbow trout Oncorhynchus mykiss: the quantitative genetics of sexual dimorphism and genotype-by-environment interactions. Aquaculture, 228: 53-68.

Keightley, P.D. 1998. Genetic basis of response to 50 generations of selection on body weight in inbred mice. Genetics, 148: 1931-1939.

Kennedy B.W. and Trus, D. 1993. Considerations of genetic connectedness between management units under an animal model. J.Anim. Sci., 71: 2341-2352.

Kennedy, B.W., van der Werf, J.H.L. and Meuwissen, T.H.E. 1993. Genetic and statistical properties of residual intake. J. Anim. Sci., 71: 3239-3250.

Kent, M. 1990. Hand-held instrument for fat/water determination in whole fish. Food Control, 1: 47-53.

Kimura, M., 1970. Stochastic processes in population genetics, with special references to distribution of gene frequencies and probability of gene fixation, p. 178-209. In: Kojima, K. (ed.) Mathematical topics in population genetics. Springer Verlag, Berlin-Heidelberg-New York.

Kerr, R.J., Goddard, M.E. and Jarvis, S.F. 1998. Maximising genetic response in tree breeding with constraints on group coancestry. Silvae Genetics, 47: 2-3.

Kincaid, H.L. 1976a. Effects of inbreeding on rainbow trout populations. Trans. Am. Fish. Soc., 105: 273-280.

Kincaid, H.L. 1976b. Inbreeding in rainbow trout *(Salmo gairdneri)*. J. Fish. Res. Board Can., 33: 2420-2426.

Kincaid, H.L., Bridges, W.R. and Von Limbach, B. 1977. Three generations of selection for growth rate in fall spawning rainbow trout. Trans. Am. Fish. Soc., 106: 621-629.

Kinghorn, B. P. 1981. Quantitative genetics in fish breeding. Ph. D. Thesis, University of Edinburgh, Edinburgh, 142 pp.

Kinghorn , B.P. 1983a. A review of quantitative genetics in fish breeding. Aquaculture, 31: 283-304.

Kinghorn, B.P. 1983b. Genetic variation in food conversion efficiency and growth in rainbow trout. Aquaculture, 32: 141-155.

Kirpichnikov, V.S. 1970. Goals and methods in carp selection. In selective breeding of carp and intensification of fish breeding in ponds, Kirpichnikov, V.S. (Editor). Israel program forscientific translations, Jerusalem, Israel.

Kirpichnikov, V.S. 1972. Methods and effectiveness of breeding the Ropshian carp. Communication I. Purpose of breeding, initial forms, and system of crosses. Sov. Genet., 8: 996-1001.

Kirpichnikov, V.S., 1981. Genetic basis of fish selection. Springer Verlag, Berlin – Heidelberg – New York. 410 pp.,

Kirpichnikov, V. S., Ilyassov, Y.I., Shart, L.A., Vikhman, A.A., Ganchenko, M.V., Ostashevsky, L.A., Smirnov, V.M., Tikhonov, G.F. and Tjurin, V.V. 1993. Selection of krasnodar common carp *(Cyprinus carpio L.)* for resistance to dropsy: principal results and prospects. Aquaculture, 111: 7-20.

Klupp, R., Heil, G and Pirchner, F. 1978. Effects of interaction between strains and environment on growth traits in rainbow trout *(Salmo gsirdneri).* Aquaculture, 14: 271-275.

Knibb, W. R. 2000. Genetic improvement of marine fish – which method for industry? Aquaculture Research, 31: 11-23.

Knibb, W.R., Gorshkova, G. and Gorshkov, S. 1997. Growth of strains of gilhead seabream Sparus aurata L. Israeli. Journal of Aquaculture bamidgeh, 49: 57-66.

Knott, S.A. and Haley, C.S. 1992. Maximum likelihood mapping of Quantitative trait loci using full-sib families. Genetics, 132(4): 1211-1222.

Kocher, T. D., Lee, W., Sobolewska, H., Penman, D. and McAndrew, B. 1997. A genetic linkage map of a cichlid fish, the tilapia *(Oreochromis niloticus)*. Genetics, 148: 1225-1232.

Kolstad, K, Heuch, P.A., Gjerde, B., Gjedrem, T. and Salte, R. 2004. Genetic variation in resistance of Atlantic salmon to the salmon louse. Aquaculture, submitted.

Komen, J. 1990. Clones of common carp, Cyprinus carpio. New perspectives in fish research. Doctoral thesis, Agricultural University Wageningen, 169 pp.

Kosambi, D. D. 1944. The estimation of map distances from recombination values. Ann. Eugenics, 12:172-175.

Kossmann, H. 1972. Untersuchungen uber die genetische Varianz der Zwischenmuskelgræten des Karpfens. Theoretical and Applied Genetics, 42: 130-135.

Lande, R. 1981. The minimum number of genes contributing to quantitative variation between and within populations. Genetics, 99: 541-553.

Lande, R. 1995. Mutation and conservation. Conserve. Biol., 9: 782-791.

Lande, R. and Thompson, R. 1990. Efficiency of marker assisted selection in the improvement of quantitative traits. Genetics, 124: 743–756.

Lander, E.S. and Botstein, D. 1989. Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics, 121: 185-199.

Langdon, C.J.,Jacobson, D.P., Evans, F. and Blouin, M.S. 2000. The molluscan brodstock program – Improving pacific oyster broodstock through genetic slection. J. Shellfish Res., 19: 616-

Lannan, J. 1972. Estimating heritability and predicting response to selection for the pacific oyster, *Crassostrea gigas*. Proc. Natl. Shellfish Assoc., 62: 62-66.

Large, R.V. 1976. The influence of reproduction rate on the efficiency of meat production in animal populations. In: D. Lister, D.N. Rodes, V.R. Fowler and M.F. Fuller (Editors), Meat Animals-Growth and Productivity. Plenum Press, New York and London, 43-55 pp.

Lee, W. 2003. Detection of QTL for salinity tolerance of Tilapia. Aquaculture Workshop, January 2003, San Diego.

Lerner I.M. 1950. Population Genetics and Animal Improvement as Illustrated by the Inheritance of Egg Production, pp. 1-74. Cambridge University Press, Cambridge.

Lewis, R.C. 1944. Selective breeding of rainbow trout at Hot Creek Hatchery. California Fish and Game, 30: 95-97.

Lie, O., Danzmann, R., Guyomar, R., Holm, L.E., Powell, R., Slettan, A. and Taggart, J. 1997. Genetic linkage maps of salmonid fishes: SALMAP. Plant and Animal Genome V, San Diego, January 1997. Abstract.

Lin, S.Y. 1940. Fish culture in the New Territories of Hong Kong. Journal of the Hong Kong Fisheries Research Station 1 (2).

Lincoln, R.F. 1981a. Sexual maturation in male triploid place (Pleuronectes platessa) and place x flounder (*Platichthys flesus)* hybrids. J. Fish Biol., 19: 415-426.

Lincoln, R.F. 1981b. Sexual maturation in female triploid place (Pleuronectes platessa) and place x flounder *(Platichthys flesus)* hybrids. J. Fish Biol., 19: 499-507.

Lincoln, R.F., Aulstad, D. and Grammeltvedt, A. 1974. Attemted triploid induction in Atlantic salmon *(Salmo salar)* using cold shocks. Aquaculture, 4: 287-297.

Lincoln, R.F. and Scott, A.P. 1983. Production of all-triploids in rainbow trout. Aquaculture, 30: 375-380.

Ling, S.W. 1977. Aquaculture in Southeast Asia: a Historical Overview. University of Washington Press, Seattle, USA. 108 pp.

Lipkin, E., Mosig, M. O., Darvasi, A., Ezra, E., Shalom, A., Friedman, A. and Soller, M. 1998. Quantitative trait locus mapping in dairy cattle by means of selective milk DNA pooling using dinucleotide microsatellite markers: analysis of milk protein percentage. Genetics, 149: 1557-1567.

Liu, S., Liu, Y., Zhang, G., Feng, H., He, X., Zhu, G. and Yang, H. 2001. The formation of tetraploid stocks of red crucian carp x common carp hybrids as an effect of interspecific hybridisation. Aquaculture, 192: 171-186.

Lodi, E. 1979. Experientia, 35: 1440-1441.

Longalong, F.M., Eknath, A.E. and Bentsen, H.B. 1999. Response to bi-directional selection for frequency of early maturing females in Nile tilapia *(Oreochromis niloticus)*. Aquaculture, 178: 13-25.

Longwell, A.C. and Stiles, S.S. 1973. Gamete cross incompatibility and inbreeding in the commercial American oyster, Crassostrea virginica Gmelin. Cytologia, 38: 521-533.

Lorenzen, N., Olesen, N.J. and Koch, C. 1999. Imunity to VHS virus in rainbow trout. Aquaculture, 172: 41-61.

Lou, Y.D. and Purdom, C.E. 1984a. Polyploidy induced by hydrostatic pressure in rainbow trout, *Salmo gairdneri Richardson.* J. Fish Biol., 25: 345-351.

Lou, Y.D. and Purdom, C.E. 1984b. Diploid gynogenesis induced by hydrostatic pressure in rainbow trout, *Salmo gairdneri Richardson*. J. Fish Biol., 24: 665-670.

Lovshin, L.L. 1980. Progress report on fisheries development in northeast Brazil. Res. Dev. Ser. No. 26. 15 pp. Culture of monosex and hybrid tilapia. FAO/CIFA Tech. Pap. 4 (Suppl. 1): 548-564.

Lund, T., Gjedrem, T., Bentsen, H.B., Eide, D.M., Larsen, H.J.S. and Røed, K.H. 1995. Genetic variation in immune parameters and associations to survival in Atlantic salmon. J. Fish Biol. 46: 748-758.

Lura, H. and Økland, F. 1994. Content of synthetic astaxanthin in escaped farmed Atlantic salmon, Salmo salar L., ascending Norwegian rivers. Fish. Manage. And Ecol., 1: 205-216.

Lura, H. and Sægrov, H. 1991. Documentation of successful spawning of escaped farmed female Atlantic salmon, *Salmo salar*, in Norwegian rivers. Aquaculture, 98: 151-159.

Lush, J.L. 1947. Family merit and individual merit as basis for selection. Part I. Am. Nat. 81: 241-261. Part II. Am. Nat., 81: 362-379.

Lush, J.L. 1949. Animal breeding plans. A book of the Iowa State College Press. 443 pp.

Lush, J.L. 1994. The genetics of populations. Iowa Agriculture and Home Economics Experiment Station College of Agriculture, Iowa State University, Ames, Iowa, 900 pp.

Lutz, C.G. 2001. Practical genetics for aquaculture. Fishing News Books, Blackwell Science, 235 pp.

Lynch, M. and Walsh, B. 1998. Genetics and analysis of quantitative traits. Sinauer Associates, Inc. 980 pp.

Macaranas, J.M., Agustin, L.Q., Ablan, M.C.A., Pante, M.J.R., Eknath, A.A. and Pullin, R.S.V. 1995. Genetic improvement of farmed tilapias: biochemical characterization of Strain differences in Nile tilapia. Aquaculture International, 3: 43-54.

MacArthur, J.W. 1949. Selection for small and large body size in the house mouse. Genetics, 34: 194-209.

Maclean, J. 1984. Tilapia, the aquatic chicken. ICLARM Newletter. ICLARM, Makati, metro Manila. 1(7).

Mahapatra, K.D., Gjerde, B., Saha, J.N., Reddy, P.V.G.K., Jana, R.K., Sahoo, M. and Rye,M. 2004. Realised genetic gain for growth in rohu *(Labeo rohita)*. Aquaculture, in manuscript.

Mair, G.C., Abucay, J.S., Beardmore, J.A. and Skibinski, D.O.F. 1995. Growth performance trials of genetically male tilapia GMT) derived from YY males in *Oreochromis niloticus L.*: On-station comparisons with mixed sex and sex reversed male populations. Aquaculture, 137: 313-322.

Mair, G.C., Abucay, J.B., Skibinski, D.O.F., Abella, T.A. and Beardmore, J.A. 1997. Genetic manupulation of sex-ratio for the large-scale production of all-male tilapia, *Oreochromis niloticus.* Canadian Journal of Fisheries and Aquatic Sciences, 54: 396-404.

Mair, G.C., Scott, A., Penman, D.J., Beardmore, J.A. and Skibinski, D.O.F. 1991. Sex determination in the genus *Oreochromis* I: Sex revesal, gynogenesis, and triploidy in *O. niloticus L*. Theor. Appl. Genet., 82: 144-152.

Malecot, G. 1948. Les mathematiques de l`heridite. Masson, Paris.

Mallet, A.L., Freeman, K.R. and Dickie, L.M. 1986. The genetics of production characters in the blue mussel *Mytilus edulis*. I. A preliminary analysis. Aquaculture, 57: 133-140.

Mallet, A.L. and Haley, L.E. 1983. Effects of inbreeding on larval and spat performance in the American oyster. Aquaculture, 33: 229-235.

Martinez, V.A., Hill, W.G. and Knott, S.A. 2002. On the use of double haploids for detecting QTL in outbred populations. Heredity, 88: 423-31.

Maynard, L.A. and Loosli, J.K. 1969. Animal nutrition. 6th edition, McGraw-Hill, New york.

McCarthy, J.C. and Siegel, P.B. 1983. A review of genetical and physiological effects of selection in meat-type poultry. Animal Breeding Abstract, 51: 87-94.

McConnell, S. K. Beynon, C., Leamon J. and Skibinski D. O. F. 2000. Microsatellite marker based genetic linkage maps of Oreochromis aureus and O. niloticus (Cichlidae): extensive linkage group segment homologies revealed. Anim. Gen., 31: 214-218.

McDonald, G.J., Danzmann, R.G. and Ferguson, M.M. 2004. Relatedness determination in the absence of pedigree information in three cultured strains of rainbow trout (Oncorhynchus mykiss). Aquaculture, 233: 65-78.

McGinnity, P., Prodöhl. P., Ferguson, A., Hynes, R., Ò Maoiléidigh, N., Baker, N., Cotter, D., O'Hea, B., Cooke, D., Rogan, G., Taggart, J., and Cross, T. 2003. Fitness reduction and potential extinction of wild populations of Atlantic salmon, *Salmo salar*, as a result of interactions with escaped farm salmon. Proc. R. Soc. Land. B

McKay. L.R., Ihssen, P.E. and Friars, G.W. 1986. Genetic parameters of growth in rainbow trout, *Salmo gairdneri,* as a function of age and maturity. Aquaculture, 58: 241-254.

Melamed, P., Gong, Z., Fletcher, G. and Hew, C.L. 2002. The potential impact of modern biotechnology on fish aquaculture. Aquaculture, 204: 255-269.

Menasveta, P. and Fast, A.W. 1998. The evolution of shrimp culture and its impact on mangroves. Infofish International, 1/98: 24-29.

Mendel, J.G. 1866. "Versuche uber Pflantzen - Hybriden". Verh. Naturforsch. Verein. Brunn, 4: 3-47.

Mesa, M.G. 1991. Variation in feeding, aggression, and position choice between hatchery and wild trout in an artificial stream. Transaction of the American Fisheries Society, 116: 574-579.

Meske, Ch. 1968. Breeding carp for reduced number of intramuscular bones, and growth of carp in aquaria. Badmidgeh, 20: 105-119.

Meuwissen, T.H.E. 1997. Maximizing the response of selection with a predefined rate of inbreeding. J. Anim. Sci., 75: 934-940.

Meuwissen, T.H.E. and Goddard, M.E. 1996. The use of marker haplotypes in animal breeding schemes. Genetics Selection Evolution, 28: 161-176.

Meuwissen, T.H.E. and Lou, Z. 1992. Computing inbreeding coefficients in large populations. Genet. Sel. Evol., 24: 305-313.

Meuwissen, T.H.E. and Sonesson, A.K. 1998. Maximizing the response to selection with a predefined rate of inbreeding: Overlapping generations. J. Anim. Sci., 76: 2575-2583.

Meuwissen, T.H.E. and Woolliams, J.A. 1994a. Effective size of livestock production to prevent a decline in fitness. Theor. Appl. Genet., 89: 1019-1026.

Meuwissen, T.H.E. and Woolliams, J.A. 1994b. Response versus risk in breeding schemes. In: Proc. of the 5th World Congress on Genetics Applied to Livestock Production, University of Guelph, Guelph, Ontario Canada, 18: 236-243.

Mitchell. G. C., Smith, M., Makower, P. J. and Bird, W. N. 1982. An economic appraisals of pig improvement in Great Britain. Genetic and production aspect. Animal Prod., 35: 215-224.

Mjølner, I.B., Refseth, U.H., Karlsen, E., Balstad, T., Jakobsen, K.S. and Hindar, K. 1997. Genetic differences between two wild and one farmed population of Atlantic salmon (Salmo salar) revealed by three classes of genetic markers. Hereditas, 127: 239-248.

Moav, R. 1976. Genetic improvement in the aquaculture industry. Pages 610-622 in: Advances in aquaculture, FAO Technical Conference on Aquaculture. Kyoto, Japan, 26 May-2 June.

Moav, R. and Finkel, A. 1975. Variability of intramuscular bones, vertebrate, ribs, dorsal fin rays and skeletal disorders in common carp. Theoretical and Applied Genetics, 46: 33-43.

Moav, R., Hulata, G. and Wohlfarth, G. 1975. Genetic differences between the chinese and European races of the common carp. 1. Analysis of genotype - environment interactions. Heredity, 34: 323-340.

Moav, R. and Wohlfarth, G.W. 1963. Breeding schemes for the genetic improvement of edible fish. Progress report. 1962. Fish Breeders Association of Israel, 44 pp.

Moav, R. and Wohlfarth, G.W. 1973. Carp breeding in Israel. In: R. Moav (Editor) Agricultural Genetics. Selected Topics. J. Wiley, New York, NY, 352 pp.

Moav; R. and Wohlfarth, G.W. 1976. Two way selection for growth rate in the common carp *(Cyprinus carpio L.)*. Genetics, 82: 83-101.

Moen, T., Høyheim, B., Munck, H. and Gomez-Raya, L. 2004a. A linkage map of Atlantic salmon (*Salmo salar*) reveals an uncommonly large difference in recombination rate between the sexes. Animal Genetics, Accepted.

Moen, T., Munck, H., Fjalestad, K.T. and Gomez-Raya, L. 2004b. A multi-stage testing strategy for detection of quantiative trait loci affecting disease resistance in Atlantic salmon. Genetics, Accepted.

Moore, S. S., Whan, V., Davis, G. P., Byrne, K., Hetzel, D. J. S. and Preston, N. 1999. The development and application of genetic markers for the Kumara prawn, *Penaues japonicus*. Aquaculture, 173: 19-32.

Moyle, P.B. 1969. Comparative behavior of young brook trout of domestic and wild origion. Prog. Fish Cult., 31: 51-59.

Mrode, R.A. 1996. Linear models for the prediction of animal breeding values. CAB International, Wallingford, Oxon OX10 8DE, UK, 187 pp.

Mustafa, A. and MacKinnon, B.M. 1999. Genetic variation in susceptibility of Atlantic salomon to the louse Caligus elongates Nordmann, 1832. Can. J. Zool., 77: 1332-1335.

Myers, J.M., Hershberger, W.K. and Iwamoto, R.N. 1986. The induction of tetraploidy in salmonids. Journal of the World Aquaculture Society, 17: 1-7.

Mørkøre, T. and Rørvik, K.A. 2001. Seasonal variation in growth, feed utilisation and product quality of farmed Atlantic salmon (Salmo salar) transferred to seawater as 0+ smolts or 1+ smolts. Aquaculture, 199: 145-157.

Naciri-Graven, Y., Martin, A.G., Baud, J.P., Renault, T. and Gerard, A. 1998. Selecting the flat oyster *Ostrea edulis (L.*) for survival when infected with the parasite Bonamia ostreae. J. Exp. Mar. Biol. Ecol.., 224: 91-107.

Nei, M. 1975. Molecular population genetics and evolution. Noth-Holland, Amsterdam.

Nell, J.A., Cox, E., Smith, I.R., and Maguire, G.B. 1994. Studies on triploid oysters in Australia.1. The farming potential of triploid Sydny rock oysters *Saccostrea commercialis* (Iredale and Roughley). Aquaculture, 126: 243-255.

Nell, A.J. and Hand, R.E. 2003 Evaluation of the progeny of second-generation Sydney rock oyster *Saccostrea glomerata* (Gold, 1850) breeding lines for resistance to QX disease Marteilia sydneyi. Aquaculture, 228: 27-35.

Nell, J.A., Smith, I.R. and Sheridan, A.K. 1999. Third generation evaluation of Sydney rock oyster Saccostrea commercialis (Iredale and Roughley) breeding lines. Aquaculture, 170: 195-203.

New, M.B. 1991. Turn of the millennium aquaculture. World Aquaculture, 22(3): 28-49.

Newkirk, G.F. 1978. Interaction of genotype and salinity in larvae of oyster crassostrea virginica. Mar. Biol. 48(3): 227-234.

Newkirk, G.F. 1980. Rewiew of the genetics and the potential for selective breeding of commercially important bivalves. Aquaculture, 19:209-228.

Newkirk, G.F. and Haley, L.E. 1983. Selection for growth rate in the European oyster, Ostrea edulis: response of second generation groups. Aquaculture, 33:149-155.

Nicholas, F.W., 1989. Incorporation of new reproductive technlogies in genetic improvement programmes. In: Evolution and Animal Breeding (Hill, W.G., MacKay, F.C., eds.). CBA International, Wallingford, pp. 203-209.

Nichols, K. M., Young, W. P., Danzmann, R. G. Robison, B. D., Rexroad, C., Noakes, M., Phillips, R. B. Bentzen, P., Spies, I., Knudsen, K., Allendorf, F. W., Cunningham, B. W., Brunelli, J., Zhang, H., Ristow, S., Drew, R., Brown, K. H. Wheeler, P. A. and Thorgaard, G. 2003. A consolidated linkage map for rainbow trout (*Oncorhynchus mykiss*) Anim. Gen. 34. 102-115.

Nilsson, J. 1992. Genetic variation in resistance of Arctic char to fungal infection. J. Aquatic Anim. Health, 4: 36-47.

Norris, A.T., Bradly, D.G. and Cunnigham, E.P. 2000. Parentage and relatedness determination in farmed Atlantic salmon (Salmo salar) using microsatellite markers. Aquaculture, 182: 73-83.

Nævdal, G., Holm, M., Møller, D. and Øvsthus, O.D. 1975. Experiments with selective breeding of Atlantic salmon. Int. Council Exp. Sea. C.M. M: 22.

O'Farrell, M.M. and Price, R.E. 1989. The occurrence of a gynandromorphic migratory trout, *Salmo trutta L*. J. Fish Biol., 34: 327.

Okamoto,N, Tayaman, T. Kawanobe, M. Fujiki, N. Yasuda, Y. and Sano, T. 1993. Resistance of a rainbow trout strain to infectious pancreatic necrosis. Aquaculture, 117: 71-76.

Olesen, I., Gjedrem, T., Bentsen, H.B., Gjerde, B. and Rye, M. 2003. Breeding programs for sustainable aquaculture. Journal of Applied Aquaculture, 13: 179-204.

Olesen, I., Gjerde, B. and Groen, Ab.F. 1999. Methodology for deriving non-market trait values in animal breeding goals for sustainable production systems. GIFT Workshop Wageningen, Session 1a Breeding and goals. Bulletin no., 23: 13-21.

Olesen, I., Groen, A.F. and Gjerde, B. 2000. Definition of animal breeding goals for sustainable production systems. Journal of Anima Science, 78: 570-582.

Osman, H. E. S. and Robertson, A. 1968. The introduction of genetic material from inferior into superior strains. Genet. Res. Camb., 12: 212-236.

Ozaki, A., Sakamoto, T., Khoo, S., Nakamura, K., Coimbra, M. R. M., Akutsu, T. and Okamoto, N. 2000. Quantitative trait loci (QTLs) associated with resistance/ susceptability to infectious pancreatic necrosis virus (IPN) in rainbow trout *(Oncorhynchus mykiss)*. Mo. Genet. Genomic, 265: 23-31.

Pante, M.J.R., Gjerde, B. and McMillan, I. 2001a. Effect of inbreeding on body weight at harvest in rainbow trout, *Oncorhynchus mykiss*. Aquaculture, 192: 201-211.

Pante, M.J.R., Gjerde, B. and McMillan, I. 2001b. Inbreeding levels in selected populations of rainbow trout, *Oncorhyncus mykiss*. Aquaculture, 192: 213-224.

Pante, M.J.R., Gjerde, B., McMillan, I. and Misztal, I., 2002. Estimation of additive and dominance genetic variances for body weight at harvest in rainbow trout, *Oncorhyncus mykiss*. Aquaculture, 204: 383392.

Pearson, K. 1903. Mathematical contributions to the theory of evolution. XI. On the influence of natural selection on the variability and correlation of organs. Philosophical Transaction of the Royal Society London Series A 200: 1-66.

Pillay, T.V.R. 1990. Aquaculture. Principles and practices. Fishing News Books, 576 pp.

Pillay, T.V.R. and Dill, W.A. 1979. Advances in aquaculture. Fishing News Books Ltd. Farnham, Surrey, England, 653 pp.

Pirchner, F. 1985. Genetic structure of populations 1. Closed populations or matings among related individuals. In World animal science. General and quantitative genetics, A. Basic information. 4. Editor A.B. Chapman, Elsevier, 227- 250.

Pondzoni, R.W., Hamzah, A.b., Saadiah, S.T. and Kamaruzzaman, N. 2003. Phenotypic and genetic parameters for live weight in two environments in a selected line of Nile tilapia in Malaysia.. International Association for Genetics in Aquaculture, Genetics in Aquaculture VIII. 9-15 November 2003: 83, Puerto Varas, Chile.

Pruginin, Y., Rothbard, S., Wohlfarth, G., Harvey, A., Moav, R. and Hulata, G. 1975. All-male broods of *Tilapia nilotica* X *T. aurea* hybrids. Aquaculture, 11, 329-334.

Purdom, C.E. 1969. Radiation-induced gynogenesis and androgenesis in fish. Heredity, London, 24: 431-444.

Purdom, C.E. 1970. Gynogensis - a rapid method for producing inbred lines of fish. Fishing News International, September, 29-32.

Purdom, C.E. 1972a. Genetics in fish farming. MAFF, Lowestoft, Lab. Leaflet, 25: 1-16.

Purdom, C.E. 1972b. Induced polyploidy in place *(Pleuronectes platessa)* and its hybrid with flounder *(Platichthys* flesus). Heredity, 29: 11-24.

Purdom, C.E. 1974. Variation in fish. In: F.R. Harden Jones (Editor), Sea fisheries Research, Elek Science, London, 347-355 pp.

Purdom, C.E. 1976. Genetic techniques in flatfish culture. J. Fish Res. Board Can., 33: 1088-1096.

Purdom, C.E. 1983. Genetic engineering by the manipulation of chromosomes. Aquaculture, 33: 287-300.

Purdom, C.E. 1993. Genetics and fish breeding. Chapman and Hall, London-Glasgow-New York- Tokyo-Melbourne-Madras. 277 pp.

Quaas, R.L. 1976. Computing the diagonal elements of a large numerator relationship matrix. Biometrics, 32: 949-953.

Quillet, E., Fosil, Land, and Chourrout, D. 1991. Production of all- triploid and all-female brown trout for aquaculture. Aquat. Living Resour./Ressour. Vivantes Aquat., 4: 27-32

Rauw, W.M., E. Kanis, E.N. Noordhuizen-Stassen, and F.J. Grommers. 1998. Undesirable side effects of selection for high production efficiency in farm animals: a review. Livest.Prod.Sci., 56: 15-33.

Reagan Jr, R.E. 1980. Heritability and genetic correlations of desirable commercial traits in channel catfish. Compl. Rep. Miss. Agric For. Exp. Stn. Publ. by Mafes-MS (USA) (June 1980) 4-5 pp.

Reagan, R.E. and Conley, C.M. 1977. Effect of egg diameter on growth of channel catfish. Prog. Fish. Cult. , 39: 133-134.

Reddy, P.V.G.K., Gjerde, B., Tripathi, S.D., Jana, R.K., Mahapatra, K.D., Gupta, S.D., Saha, J.N. Lenka, S., Sahu, M., Govindassamy, P., Rye, M. and Gjedrem, T. 2002. Growth and survival of six stocks of rohu (*Labeo rihita*) in mono- and polyculture production systems. Aquaculture, 203: 239-250.

Reed, D.H. and Frankham R., 2001**.** How Closely Correlated Are Molecular and Quantitative Measures of Genetic Variation? A Meta-Analysis**.** Evolution, 55(6):1095-1103.

Refstie, T. 1981. Tetraploid rainbow trout produced by cytochalasin B. Aquaculture, 25: 51-58.

Refstie, T. 1982. Preliminary results: differences between rainbow trout families in resistance against vibriosis and stress: In: W.B van Muiswinkel (Editor), Developmental and ComparativeImmunology, suppl. 2. Pergamon Press, New York, NY, 205-209, pp.

Refstie, T. 1983a. Hybrids between salmonid species. Growth rate and survival in seawater. Aquaculture, 33: 281-285.

Refstie, T. 1983b. Indiction of diploid gynogenesis in Atlantic salmon and rainbow trout using irradiated sperm and heat shock. Can. J. Zool., 61: 2411-2416.

Refstie, T. 1990. Application of breeding schemes. Aquaculture, 85: 163-169.

Refstie, T. and Aulstad, D. 1975. Tagging experiments with salmonids. Aquaculture, 5: 367-374.

Refstie, T and Steine, T. 1978. Selection experiments with salmon. III. Genetic and environmental sources of variation in length and weight of Atlantic salmon in the freshwater phase. Aquaculture, 14: 221-234.

Refstie, T., Steine, T. and Gjedrem, T. 1977. Selection experiments with salmon. II. Proportions 0f Atlantic salmon smoltifying at one year of age. Aquaculture, 10: 231- 242.

Refstie, T., Stoss, J. and Donaldson, E.M. 1982. Production of all female coho salmon *(Oncorhynchus kisutch)* by diploid gynogenesis using irradiated sperm and cold shock. Aquaculture, 29: 67-82.

Refstie, T., Vassvik, V. and Gjedrem, T. 1977. Induction of polyploidy in salmonids by Cytochalasin B. Aquaculture, 10: 65-74.

Refstie, T., Mørkøre, T., Johansen, H. and Gjøen, H.M. 1999. Texture. AKVAFORSK Report no. 12/99, 19 pp.

Rezk, M.A., Smitherman, R.O., Williams, J.C., Nichols, A., Kucuktas, H. and Dunham, R.A. 2003. Response to three generations of selection for increased body weight in channel catfish, *Ictalurus punctatus*, grown in earthen ponds. Aquaculture, 228: 69-79.

Robertson, A. 1955. Selection response and the properties of genetic variation. Cold. Spr. Harb. Symp. Quant. Biol., 20: 166-177.

Robertson, A. and Lerner, I.M. 1949. The heritability of all-or-non traits: viability of poultry. Genetics, 34: 395.

Robinson, N. 2002. Techniques for advansed genetic improvement of fish and shellfish in aquaculture. Report on an NRE scientific exchange visit to aquaculture and marine science institute in Scotland and Norway, September 2002. Victorian Institute of Animal Science, Attwood VIC 3049, Australia, 35pp.

Robison, O.W. and Luempert, L.G. 1984. Genetic variation in weight and survival of brook trout *(Salvelinius fontinalis).* Aquaculture, 38: 155-170.

Romashov, D.D., Belyaeva, V.N., Golovinskaia, K.A. and Prokofieva-Bel'govskaya, A.A. 1961. Radiation disease in fish. Radiatsionnaya Genetika, 161: 247-266.

Ruzzante, D.E. 1994. Domestication effects on aggressive and schooling behavior in fish. Aquaculture, 120: 1-24.

Ruzzante, D.E. and Doyle, R.W. 1991. Rapid behavioral changes in medaka, *Oryzias latipes*, during selection for competitive and noncompetitive growth. Evolution, 45: 1936-1946.

Rye, M. 1991. Prediction of carcass composition in Atlantic salmon by computerised tomography. Aquaculture, 99: 35-48.

Rye, M., Johansen, H., Suarez, J.A., Angarita, M.R. and Gitterle, T. 1999. Selective breeding of *Litopenaeus vannamei* in Columbia. Batch 1, 1998. Technical Report, December 1999. Institute of Aquaculture Research Ltd. Norway and Corporation Centro de Investigaciona de Columbia (CENIACUA), Columbia, 21 pp.

Rye, M., Bæverfjord, G. and Jopson, N. 1994. Datomografi avslører feittfordelinga hos laks. (Computerized tomographi reveal distribution of fat in salmon). Norsk Fiskeoppdrett, 11A-94: 37-39.

Rye, M. and Eknath, A.E. 1999. Genetic improvement of tilapia through selective breeding – Experience from Asia. European Aquaculture Society, Special Publication. No 27; June 1999: 207-208.

Rye, M. and Gjerde, B. 1996. Phenotypic and genetic parameters of composition traits and flesh colour in Atlantic salmon. Aquaculture Research, 27: 121-133.

Rye, M., Lillevik, K.M. and Gjerde, B. 1990. Survival in early life of Atlantic salmon and rainbow trout: estimates of heritabilities and genetic correlations. Aquaculture, 89, 209-216.

Rye, M and Mao, I.L. 1998. Nonadditive genetic effects and inbreeding depression for body weight in Atlantic salmon *(Salmo salar L.).* Lives. Prod. Sci., 57: 15-22.

Rye, M. and Refstie, T. 1995. Phenotypic and genetic parameters of body size traits in Atlantic salmon, *Salmo Salar L*. Aquaculture research, 26: 875-885.

Rye, M., Storebakken, T., Gjerde, B. and Ulla, O. 1994. Efficient selection for colour in Atlantic salmon (Effektiv seleksjon for innfarging hos laks). Norsk Fiskeoppdrett, 11A: 22-24.

Ryman, N. 1970. A genetic analysis of recapture frequencies of released young of salmon *(Salmo salarL.)*. Hereditas, 65: 159-160.

Ryman, N. 1981 (ed.). Fish Gene Pools. Preservation of Genetic Resources in Relation to Wild Fish Stocks. Ecological Bulletins (Stockholm), 34:1-111.

Røed, K.H., Brun, E., Larsen, H.J. and Refstie, T. 1990. The genetic influence on serum haemolyticactivity in rainbow trout. Aquaculture, 85: 109-117.

Røed, K.H., Fjalestad, K., Larsen, H.J. and Midthjell, L. 1992. Genetic variation in haemolytic activity in Atlantic salmon *(Salmo salar L.).* J. Fish Biol., 40: 739-750.

Røed, K.H., Fjalestad, K.T. and Strømsheim, A. 1993. Genetic variation in lysozyme activity and pontaneous haemolytic activity in Atlantic salmon *(Salmo salar).* Aquaculture, 114: 19-31.

Sakamoto, T., Danzmann, R., Okamoto, N., Ferguson, M. and Ihssen, P. 1999. Linkage analysis of quantiative trait loci associated with spawning time in rainbow trout *(Oncorhynchus mykiss)*. Aquaculture, 173: 33-43.

Salte, R., Galli, A., Falaschi, U., Fjalestad, K.T., and Aleandri, R. 2004. A protocol for the on-site use of frozen milt from rainbow trout (Oncorhynchus mykiss Walbaum) applied to the production of progeny groups: comparing males from different populations. Aquaculture, 231: 337-345.

Salte, R., Gjøen, H.M., Norberg, K. and Gjedrem, T. 1993. Plasma protein levels as potential marker traits for resistance to furunculosis. J. Fish Diseases, 16: 561-568.

Sanders, G.P. 1986. Biology the science of life. Scott, Foresman and Company. Glenviw, Illinois; London, England, 1217 pp.

Sarmasik A., Warr, G. and Chen, T.T. 2002. Production of transgenic medaka with increased resistance to bacterial pathogens. Mar. Biotechnol., 4, 310-322.

Schaperclaus, W. 1961. Lehrbuch der Tierchwirtschaft. Berlin, P. Parey. 582 pp. Purdom, C.E. 1976. Genetic techniques in flatfish culture. J. Fish Res. Board Can., 33: 1088-1096.

Schaperclaus, W. 1962. Trate de Pisciculture en Etang. Vigot Freres, Paris, 208-227 pp.

Schultz, R, J. 1971. Am. Zool. 11: 351-360.

Scott, A.P. and Baynes, S.M. 1980. A review of the biology, handling and storage of salmonid spermatozoa. J. Fish Biol., 17: 707-739.

Scott, A.G., Penman, D.J., Beardmore, J.A. and Skibinski, D.O.F. 1989. The YY supermale in *Oreochromis niloticus (L.)* and its potential in aquaculture. Aquaculture, 78: 237-251.

Searle, S.R. 1971. Linear Models. John Wiley and Sons, 532 pp.

Searle, S. R. and Henderson, C.R. 1961. Computing procedures for estimating components of variance in the two-way classification, mixed model.. Biometrics, 17: 607-616.

Sehested E, and Mao I.L. 1992. Responses to selection on genotypic or phenotypic values in the presence of genes with major effects. Theoretical and Applied Genetics, 85: 403-406.

Sengbusch, R. 1963. Fische "ohne Græte". Der Zuchter, 33: 284-286.

Sengbusch, R. 1967. Eine Schnellbestimmungsmethode der Zwischenmuskelgræten bei Karpfen zur auslese von "grætenfreien" Mutanten (mit Røntgen-Fernsehkamera und Bildschirmgeræt). Der Zuchter, 37: 275-276.

Sengbusch, R. and Meske, C. 1967. Auf dem Wege zum grætenlosen Karpfen. Der Zuchter, 37: 271-274.

Shadidi, F. and Botta, J.R. 1994. Seafoods. Chemistry, processing technology and quality. Blackie Academic and Professional, London, Glasgow, Weinheim, New York, Tokyo, Melbourne, Madras.

Shultz, F.T. 1986. Developing a commercial breeding program. Aquaculture, 57: 65-76.

Siitonen, L. and Gall, G.A.E. 1989. Response to selection for early spawn date in rainbow trout, *Salmo gairdneri.* Aquaculture, 78: 153-161.

Silverstein, J. 2003. Use of residual feed intake to compare feed efficiencies in strains of rainbow trout, *Onchorhynchus mykiss*. Abstract, Genetics in Aquaculture VIII, 9-15 Novenber, 2003, Puerto Varas – Chile, 77 pp.

Silverstein, J.T., Bosworth, B.G., Waldbieser, G.C. and Wolters, W.R. 2001. Feed intake in channel catfish: is there a genetic component? Aquaculture Research, 32: 199-205.

Singer, M. and Berg, P. 1991. Genes & Genomes. University Science Books, Mill Valley, CA.

Singer, A., Perlman, H., Yan, Y., Walker, C., Corley-Smith, G., Brandhorst, B. and Postlethwait, J. 2002. Sex specific recombination rates in zebrafish *(Danio rerio).* Genetics, 160: 649-657.

Siraj, S.S., Smitherman, R.O., Castillo-Gallusser, S. and Dunham, R.A. 1983. Reproductive traits for three year classes of Tilapia niloticus and maternal effects on their progeny. In International Symposium on Tilapia in Aquaculture, L. Fishelson and Yarson (Editors). Tel Aviv, Israel.

Skaala, Ø. and Hindar, K. 1998. Genetic changes in the River Vosso salmon stock followinga collapse in the spawning population and invation of farmed salmon, p. 29 in: Youngson, A.F., Hansen, L.P. and Windsor, M.L. (eds.) Interactions between salmon culture and wild stocks of Atlantic salmon. The scientific and management issues. Norsk institutt for naturforskning, Trondheim.

Skaala, Ø., Taggart, J.B., Nævdal, G., Gunnes, K., Jørstad, K.E. and Karlsen, T. 1999. Genetic variability of farmed Atlantic salmon strains and wild stocks. Poster, AquaNOR, 1999.

Skjervold, H. 1982. Die Bildung einer synthetischen Rasse. Archiv fur Tierzucht, 25: 1-12.

Smiesek, J. 1979. Hybridization of carp of the Vodnany and Hungarian lines. Bull. VÜRH Vodnany, 15(1): 3-12.

Smith, H. 1936. A discriminant function for plant selection. Ann. Eug., 7: 240-250.

Smith, V.K. 1993. Nonmarket valuation of environmental resources: An interpretive apprisal. Land Econ., 69: 1-26.

Snedecor, G.W. 1957. Statistical methods. Ames, Iowa, Iowa State College Press, 534 pp.

Snedecor, G.W. and Cochran, W.G. 1980. Statistical methods. (7th edn). Iowa State University, Ames, Iowa, USA, 507 pp.

Sneed, K.E. 1971. Some current North American work in hybridization and selection of cultured fishes. In: FAO Seminar/Study tour in the USSR on genetic selection and hybridization of cultivated fishes. Rep. FAO/UNDP- (TA) 2926, 143-150 pp.

Sokal, R.R. and Rolph, F.J. 1981. Biometry. The principles and practice of statistics in biological research, 2nd Edition, W.H. Freeman and Company, New York, 859 pp.

Solazzi, M.F. 1977. Effects of inbreeding coho salmon *(Oncorhynchus kisutch).* An abstract of a thesis submitted to Oregon State University, 26 pp.

Sonesson, A.K., Gjerde B. and Meuwissen, T.H.E. 2004. Truncation selection for BLUP-EBV and phenotypic values in fish breeding schemes. AKVAFORSK, manuscript.

Sonesson, A.K. and Meuwissen, T.H.E. 2000. Mating schemes for optimum contribution selection with constrained rates of inbreeding. Genet. Sel. Evol., 32: 231-248.

Sorensen, D.A. and Kennedy, B.W. 1984a. Estimation of respose to selection using least squares and mixed model nethodology. J. Anim. Sci.,58: 1097-1106.

Sorensen, D.A. and Kennedy, B.W. 1984b. Estimation of genetic variances from unselected and selected populations. J. Anim. Sci., 59: 1213-1223.

Sonesson, A.K. and Meuwissen, T.H.E. (2000) Mating schemes for optimum contribution selection with constrained rates of inbreeding. Genet. Sel. Evol., 32: 231-248.

Sorensen, D., Wang, C.S., Jensen, J. and Gianola, D. 1994. Bayesian analysis of genetic changes due to selection using Gibbs sampling. Genet. Sel. Evol., 26: 333-360.

Stabell, O.B. 1984. Homing and olfaction in salmonids: a critical review with with special references to the Atlantic salmon. Biol. Rev., 59: 333-388.

Standal, M. and Gjerde, B. 1987. Genetic variation in survival of Atlantic salmon during the sea- rearing period. Aquaculture, 66: 197-207.

Stanley, J.G. 1976. Production of hybrid, androgenetic and gynogenetic grass carp and carp. Transactions of the American Fisheries Society, 105: 10-16.

Stanley, J.G., Hidu, H. and Allen, S.K. 1984. Growth of American oysters increased by polyploidy induced by blocking meiosis I but not meiosis II. Aquaculture, 37: 147-155.

Steine, T 1980. Fiften years' experience with a co-operative sheep breeding scheme in Norway. Proc. World Congres in Sheep and Cattle Breeding, II: 145-148.

Storebakken, T., Austreng, E. and Steenberg, K. 1981. A method for determination of feed intake in salmonids using radioactive isotopes. Aquaculture, 24: 133-142.

Stoss, J. 1979. Spermakonservirung bei der Regenbogenforelle, (Salmo gairdneri). Dissertation, Institut fur Tierzucht und Haustiergenetik der Georg-August-Univeristat Gøttingen, Gøttingen, Germany, 122 pp.

Stoss, J. 1983. Fish gamete preservation and spermatozoan physiology. Fish Physiology, Vol. IXB: 305-349. Academic press, Inc. IBSN 0-12-350429-5.

Stoss, J. and Refstie, T. 1983. Short-term storage and cryopreservation of milt from Atlantic salmon and sea trout. Aquaculture, 30: 229-236.

Streelman, J.T. and Kocher, T.D. 2002. Microsatellite variation associated with prolactin expression and growth of salt-challenged tilapia. Physiol. Genomics, 9: 1-4.

Strømsheim, A., Eide, D.M., Hofgaard, P.O., Larsen, H.J.S. Refstie, T. and Røed, K.H. 1994a. Genetic variation in the humoral immune response against Vibrio salmonicida and in antibody titre against Vibrio anguillarum and total IgM in Atlantic salmon (Salmo salar). Veterinary Immunology and Immunopathology, 44: 85-95.

Strømsheim, A., Eide, D.M. Fjalestad, K.T. Larsen, H.J.S. and Røed, K.H. 1994b. Genetic variation in the humoral immune response in Atlantic salmon (Salmo salar) against Aeromonas salmonicida A-layer. Veterinary Immunology and Immunopathology, 41: 341-352.

Ståhl, G. and Hindar, K. 1988. Genetic structure of Norwegian salmon: Status and perspectives. Rapport fra Fiskeriforskningen. No. 1 (In Noewegian).

Su, Guo-Sheng, Liljedahl, L.E. and Gall, G.A.E. 1996. Effect of inbreeding and growth and reproductive traits in rainbow trout *(Onchorhyncus mykis).* Aquaculture, 142: 139-148.

Suarez, J.A., Gitterle, T., Angarita, M.R. and Rye, M. 1999. Genetic parameters for harvest weight and pond survival in *Litopenaeus vannamei*. European Aquaculture Society, Special Publication. No 27; June 1999: 232-233.

Sun, X. and Liang, L. 2003. A genetic linkage map of common carp *(Cyprinus carpio L.)* and mapping of a loci associated with cold tolerance. Aquaculture, in press.

Supan, J.E., Allen, Jr. S. K. and Wilson, C.A. 2000. Tetraploid eastern oysters: An ardous effort. J. Shellfish Res., 19: 655.

Sutherland. T.M. 1965. The correlation between feed efficiency and rate of gain, a ratio and its denominator. Biometrics, 21: 739-749.

Svardson, G. 1945. Chromosome studies of Salmonidae. Meddn St. Unders.- Förs Ans. Söttvatt. Fisk., 23: 1-. 151.

Swain, D.P. and Riddell, B.E. 1991. Domestication and agonistc behaviour in coho salmon: reply to Ruzzante. Can. J. Fish. Aquat. Sci., 48: 520-522.

Sylven, S., Rye, M. and Simianer, H. 1991. Interaction of genotype with production system for slaughter weight in rainbow trout *(Oncorhynchus mykiss)*. Livestock Production Science, 28: 253-263.

Sørensen P. 1986. Study of the effect of selection for growth in meat type chickens. (Studium af effekten af selektion for vækst hos slagtekyllinger). Report no. 612 from the National Institute of Animal Science, Denmark, 31-35 pp.

Talbot, C. and Higgins, P.J. 1983. A radiographic method for feeding studies on fish using metallic iron powder as a marker. J. Fish Biol., 23: 211-220.

Tanck, M.W.T., Vermeulen, C.J., Bovenhuis, H and Komen, J. 2001. Heredity of stress-related cortisol response in androgenetic common carp *(Cyprinus carpio L.)*. Aquaculture, 199: 283-294.

Tave, D. 1986a. Genetics for fish hatchery managers. AVI publishing Company, Inc. Westport, Connecticut.

Tave, D. 1986b. A Quantitative Genetic Analysis of 19 Phenotypes in *Tilapia nilotica*. Copeia, 3: 672-679.

Tave, D., Rezk, M. Smitherman, R.O. 1989. Genetics of body color in Tilapia mossambica. J. World Aquacul. Soc. 20, 214-222.

Tave, D. and Smitherman, R.O. 1980. Predicted response to selection for early growth in Tilapia nilitica. Trans. Am. Fish. Soc., 109: 439-445.

Tayamen, M. M., Reyes, R. A., Danting, Ma., D., Mendoza, A. M., Marquez, E. B., Salguet, A. C., Gonzales, R., C. Abella, T. A. and Vera-Cruz, E. M. 2002. Tilapia broodstock development for saline waters in the Phillipines. Naga, The ICLARM Quarterly, 25(1): 32-26.

Teichert-Coddington, D. 1983. divergent selection for prematuration body weight in tilapia niloticus. M.S. Thesis, Auburn University, Auburn, AL.

Teichert-Coddington, D.R. and Smitterman, R.O. 1988. Lack of response by tilapia nilotica to mass selection for rapid early growth. Trans. Am. Fish. Soc., 117: 297-300.

Terjesen, B.F., Barr, Y., Lein, I. and Gjerde, B. 2004. Development of a breeding nucleus production protocol for the Atlantic cod (*Gadus morhua* L.). I: Model development and production system dimensions. In manuscript.

The Tecnical Committee for tilapia. 1985. The Philippines recommends for tilapia. PCARRD Tech. Bull. Series No. 15-A. Philippines Council for Agriculture and Resources Research and development. Los Banos, Laguna.

Thodesen, J. 1999a. Selection for improved feed conversion in salmon (Avl for bedre fôrutnytting hos laks). Norsk Fiskeoppdrett, March 1999, No. 5A, 20-21.

Thodesen, J. 1999b. Selection for improved feed utilization in Atlantic salmon. Doctor Scientiarum Thesis, Agricultural University of Norway, 1999, 108 pp.

Thodesen, J., Gjerde, B., Grisdale-Helland, B. and Storebakken, T. 2001. Genetic variation in feed intake, growth and feed utilization in Atlantic salmon (Salmo salar). Aquaculture, 194: 273-281.

Thodesen, J., Grisdale-Helland, B., Helland, S.J. and Gjerde, B. 1999. Feed intake, growth and feed utilization of offspring from wild and selected Atlantic salmon (*Salmo salar*). Aquaculture, 180: 237-246.

Thorgaard, G.H. 1992. Application of genetic technologies to rainbow trout. Aquaculture, 100: 85-97

Thorgaard, G.H., Bailey, G.S., Williams, D., Buhler, D.R., Kaattari, S.L., Ristow, S.S., Hansen, J.D., Winton, J.R., Bartholomew, J.L., Nagler, J.J., Walsh, P.J., Vijayan, M.M., Devlin, R.H., Hardy, R.W., Overturf, K.E., Young, W.P., Robison, B.D., Rexroad, C. and Palti, Y. 2002. Status and opportunities for genomics research with rainbow trout. Comp. Biochem. and Phys. (Part B), 133: 609-646.

Thorgaard; G.H. and Gall, G.H. 1979. Adult triploids in a rainbow trout family. Genetics, 93: 961-973.

Thorgaard, G.H., Jazwin, M.E. and Stier, A.R. 1981. Polyploidy induced by heat shock in rainbow trout. Trans. Am. Fish. Soc., 110: 546-550.

Thorgaard, G.H., Rabinovitch, P.S., Shen, M.W., Gall, G.A.E., Propp, J. and Utter, F.M. 1982. Triploid rainbow trout identified by flow cytometry. Aquaculture, 29: 305- 309.

Thorgaard, G.H., Scheerer, P.D., Hershberger, W.K. and Myers, J.M. 1990. Angrogenetic rainbow trout produced using sperm from tetraploid males show improved survival. Aquaculture, 85: 215-221.

Toro, M.A. 1998. Selection of grandparental combinations as a procedure designed to make use of dominance genetic effects. *Genet. Sel. Evol*., 30: 339-349.

Toro, J.E., Aguila, P and Vergara, A.M. 1996. Spatial variation in response to selection for live weight and shell length from data on individually tagged Chilian native oysters (Ostrea chilensis Philippi, 1845). Aquaculture, 146: 27-36.

Toro, J.E. and Newkirk, G.F. 1990. Divergent selection for growth rate in the European oyster Ostrea edulis: response to selection and estimate of genetic parameters. Mar. Ecol. Prog. Ser., 62(3): 219-227.
Toro, J.E. and Newkirk, G.F. 1991. Response to artificial selection and realized heritability estimate for shell height in the Chilian oyster Ostrea chilensis. Aquatic Living Resources, 4(2): 101-108.
Toro, M.A., Nieto, B. and Salgado, C. 1988. A note on minimization of inbreeding in small scale selection programmes. Livest. Prod. Sci., 20: 317-323.
Toro, M. and Perez-Enciso, M. 1990. Optimization os selection response under restricted inbreeding. Genet. Sel. Evol., 22: 93-107.
Traxler, G.S., Anderson, E., LaPetra, S.E., Richard, J., Shewmaker, B. and Kurath, G. 1999. Dis. Aquat. Org., 38: 183-190.
Trong, T.Q. 2004. Walk back selection. Msc. Thesis, Agricultural University of Norway, 1430 Aas, Norway, 36pp.
Uraiwan, S. and Doyle, R.W. 1986. Replicate variance and the choice of selection procedures for tilapia *(Oreochromis niloticus)* stock improvement in Thailand. Aquaculture, 57: 93-98.
Vangen, O. 1984. Future breeding program in pigs in a situation with artificial insemination. (Framtidig avlsopplegg på svin i en KS-situasjon). Aktuelt fra Statens fagtjeneste for landbruket, 1: 300-306.
Vangen, O. and Kolstad, N. 1986. Genetic control of growth composition, appetite and feed utilization in pigs and in poultry. I.: Proc. 3rd World Congr. Genetics Applied to Livestock Prod. University of Nebraska, Lincoln, Nebraska, USA, 11: 367-380.
Van Bebber, J. and Meyer, J.T. 1994. Selection for feed efficiency in broilers: A Comparison of residual feed intake with feed conversion ratio. In: Proc. 5th World Congr. Genetics Applied to Livestock Prod. University of Guelph, Ontario, Canada, 20: 53-56.
Van Vleck, L.D., Pollak, E.J. and Oltenacu, E.A.B. 1987. Genetics for the Animal Sciences. W.H. Freeman and Company, New York, 391 pp.
Varadaraj, K. and Pandian, T.J. 1989. First report on production of supermale tilapia by integrating endocrine sex reversal with gyogenetic technique. Curr. Sci., 58: 343-441.
Vignal, A., Milan. A., SanCristobal, M. and Eggen, A. 2002. A review on SNP and other types of molecular markers and their use in animal genetics. Gen. Sel. Evol., 34: 275-305.
Villanueva, B., Verspoor, E. and Visser, P.M. 2002. Parental assignment in fish using mivrosatelilite genetic markers with finite numbers of parents and offspring. Animal Genetics, 33: 33-41.
Villanueva, B., and Woolliams, J.A. 1997. Optimization of breeding programmes under index selection and constrained inbreeding. Genet. Res. Camb., 69: 145-158.
Villanueva, B., Woolliams, J.A. and Gjerde, B. 1996. Optimum designs for breeding programmes under mass selection with an application in fish breeding. Annimal Science, 63: 563-576.
Villanueva, B., Woolliams, J.A. and Simm, G. 1994. Strategies for controlling rates of inbreeding in MOET nucleus schemes for beef cattle. Genet. Sel. Evol., 26: 517-535.
Von Limbach, B. 1970. Breeding rainbow trout. Pages 153-158 in U.S. Bureau of Sport Fisheries and Wildlife Annual Report, Devision of Fisheriy Research, 318pp.
Wallace, R.A., King, J.L. and Sanders, G.P. 1986. Biology the science of life. Scott, Foresman and Company. Glenviw, Illinois, London, England, 1217 pp.
Watson, J. and Crick, F.H.C. 1953. Molecular structure of nucleic asids. Nature, 171: 737-738.
Weinberg, W. 1908. Uber den Nachweis der Vererbung beim Menschen. Jahreshelfts Ver. vaterl. Naturf. Wurttemberg, 64: 369-382.
Weller, J.I. 2001. Quantitative trait loci analysis in animals. CABI Publishing. 287pp.WorldFish Cebter 2002. INGA Technical Report (July 2001-June 2002). Submitted to the Norwegian Agency for Development Ciooperation.
Wild, V., Simianer, H., Gjøen, H.M. and Gjerde, B. 1994. Genetic parameters and genotype – environment interaction for early sexual maturity in Atlantic salmon *(Salmo salar).* Aquaculture, 128: 51-65.
Wilson, K., Li, Y., Whan, V., Lehnert, S., Byrne, K., Moore, S., Pongsomboon, S., Tassanakajon, A., Rosenborg, G., Ballment, E., Fayazi, Z., Suan, J., Kemay, M. and Benzieg, T. 2002. Genetic mapping of the black tiger shrimp *Penaeus monodon* with amplified fragment length polymorphism. Aquaculture, 204: 297-309.
Winkelman, A.M. and Peterson, R.G. 1994. Genetic parameters (heritabilities, dominance ratios and geneticcorrelations) for body weight and length of chinook salmon after 9 and 22 months of saltwater rearing. Aquaculture, 125: 31-36.
Wohlfarth, G.W. 1983. Genetics of fish: Application to warm water fishes. Aquaculture, 33: 373-381.
Wohlfarth, G.W. 1993. Heterosis for growth rate in common carp. Aquaculture, 113: 31-46.

Wohlfarth, G., Moav, R. and Hulata, G. 1983. A genotype - environment interaction for growth rate in common carp, growing in intensive manured ponds. Aquaculture, 33: 187-195.

Woods, I.G., Kelly, P.D., Chu, F., Ngo-Hazelett, P., Yan, Y.L. Huang, H., Postlethwait, J.H. and Talbot, W.S. 2000. A comparative map of the zebrafish genome. Genome Research, 10: 1903-1914.

Woolliams, J.A. 1989. Modifications to MOET nucleus breeding schemes to improve rates of genetic progress and decrease rates of inbreeding in dairy cattle. Anim. Prod., 49: 1-14.

Woolliams, J.A., Pong-Wong, R. and Villanueva, B. 2002. Proceedings 7th World Congress on Genetics Applied to Livestock Production, Communcation, 23-32.

WorldFish Center, 2002. Jalan Batu Maung, Batu Maung, Penang, Malaysia.

Wray, N. and Hill, W.G. 1989. Asymptotic rate of responses from index selection. Anim. Prod., 49: 217-227.

Wray, N.R., Woolliams, J.A. and Thompson, R. 1994. Prediction of rates of inbreeding in populations undergoing index selection. Theor. Appl. Genet., 87: 878-892.

Wright, S., 1931. Evolution in Mendelan population. Genetics, 16: 97-159.

Wright, S. 1934. The method of path coefficients. Annals of Mathematical Statistics, 5: 161-5.

Yamamoto, T. 1969a. Inheritance of albinism in the medaka, Oryzias latipes, with special reference to gene interaction. Genetics, 62: 797-809.

Yamamoto, T. 1969b. Sex differentiation, In Fish Physiology by Hoar, W.S. and Randall, D.J., Vol. III: 117-175. Academic Press, New York and London.

Yamazaki, F. 1983. Sex control and manipulation in fish. Aquaculture, 33: 329-354.

Yang, H., Guo, X., Chen, Z. and Wang, Y. 1999. Tetraploid induction by inhibiting mitosis 1 in scallop Chlamus farreri.Chin. J. Oceanol. Limnol., 17: 350-358.

Yang, H., Wang, R., Guo, X. and Yu, Z. 1997. Tetraploid induction by blocking polar body 1 and mitosis 1 infertilized eggs of the scalop Chlamys (Azumapecten) farreri with cytochalasin B. J. Ocean Univ. Qingdao/Qingdao Haiyang Daxue Xuebao., 27: 166-172.

Young, W.P., Wheeler, P.A., Coryell, V.H., Keim P. and Thorgaard, G.H. 1998. A detailed genetic linkage map of rainbow trout produced using doubled haploids. Genetics, 148: 839-50.

Zohar, Y., Abraham, M. and Gordin, H. 1978. The gonadal cycle of the captivity-reared hermaphroditic teleost *Sparus aurata (L.)* during the first two years of life. Annales de Biologie Animale, iochimie Bet Biophsique, 18: 877-882.

Åsgård, T. and Hillestad, M. 1998. Eco-friendly aquafeeds and feeding. Safe feed safe food. Symposium Victam 98. Utrecht, The Netherlands, 13-14 May.

INDEX

Adaptation, 3, 31
Additive, 35, 121, 175
    effect, 35, 54, 121
Accuracy, 161, 176, 198, 208, 268
Allele, 23
    frequency, 23
    fixation, 39,43, 83
    maternal, 73
    paternal, 73
Amino acids, 14
Ancestor, 78
    common, 74, 78
Androgenesis, 287, 291, 327
Animal model, 214
ANOVA, 124

Base population, 174 , 260
    synthetic, 260
Behaviour, 3, 5, 259
Biotechnology, 301
Bisexual, 113, 114
BLUE, 217
BLUP, 185, 200, 214, 217, 248
Breeding
    goal, 145, 198, 253, 271, 283
    multipliers, 193, 276
    nucleus, 192, 281
    program, 6, 87, 145, 174, 251
    pure, 157, 253
    strategies, 145, 261
    station, 253, 279, 281
Breeding value, 89, 163, 178, 272
    estimated, 202
    prediction, 197, 202
    true, 198, 202
Broodstock, 251, 273, 283, 313
Bulmer effect, 101, 175

Cell, 15, 17, 19, 287
Chromosome, 10, 18, 252, 287, 303
    engineering, 287
    manipulation, 324
Challenge test, 257, 269

Coefficient
    coancestry, 73
    inbreeding, 73
    regression, 60
    relationship, 76
    variation, 46, 47
Combining ability, 150
Competition, 109, 265
Connectedness, 189
Control populations, 243, 273
Correction factor, 56
Correlation, 59
    environmental, 63, 136
    genetic, 106, 136, 235, 256, 274
    intraclass, 171
    phenotypic, 61, 136
Covariance, 60, 121
Crossbreeding, 26, 150, 261, 323
    backcross, 153
    diallel, 153
    four-way, 153
    species, 155
    three-way, 153
Cryopresevation, 119, 188

Disease resistance, 97, 257, 269
    furunculosis, 102, 257, 322
    IPN, 102
    selection, 97, 257
    Taura syndrom, 103, 258
Dissemination, 276, 281
Distribution, 47
    binomial, 38, 42
    discontinuous, 51
    non-normal, 49
    normal, 36, 47, 48, 92, 173
    skewed, 50, 110, 177
DNA, 9, 23, 301, 316, 323
    base pairs, 11, 14
    fingerprinting, 185, 306
    markers, 190, 302, 316
    polymorphism, 35
    replication, 14
Domestication, 3, 36, 109, 167, 251
Dominance, 21, 28, 43, 198

effect, 54, 121

Economic value, 166, 254, 283
Effective number, 25, 146
population size, 26, 75, 146
Effect, fixed, 200
Egg, 14, 115, 117
Environment, 272
change, 243
common, 133, 162
covariance, 122
fixed effect, 130, 230
permanent, 55, 129
random, 55, 230
systematic, 55
Epistasis, 43, 198
effect, 54, 121
Equilibrium, 5, 23, 25, 35, 232
Escapees, 324, 325, 330
Etics, 253

FCR, 95, 97, 255, 268
Fertility, 115, 252, 259, 282
Fertilization, 118, 252, 283, 287
Fitness, 84, 89, 109, 150, 274, 320, 332
Fixation, 36, 42, 107, 323
Founder effect, 30

Gene, 17, 35, 197
frequency, 24
how many, 40, 109
mapping, 21, 303
pleiotropic, 102
recessive, 27
Generation interval, 29, 93, 262, 326
Genetic
adaptation, 20, 320
change, 243, 273, 327
covariance, 122, 123, 206
distance, 31, 305, 308
diversity, 319, 320
drift, 25, 29, 31, 81
effect, 35
gain, 41, 93, 116, 184, 198, 243
linkage, 25
maps, 303
molecular, 9
parameters, 121
Genotype, 21
aggregate, 254
frequency, 23
allelic, 35
recessive, 27
GIFT tilpia, 110, 153, 260
GLS, 202, 217
Gynogenese, 113, 288, 323

Helix, double, 10
Heritability, 64, 66, 102, 132, 166, 170
dam estimate, 66, 133
family means, 133, 204
realized, 67
regression, 132
sire estimate, 66, 133
Hermaphrodite, 113, 114, 115
Heterosis, 26, 150, 156, 261
Hetrozygote, 21, 24, 28, 152
Heterozygosity, 21, ,26, 75, 83
Homozygote, 21, 24, 28, 291
Hormone treatment, 115, 118, 271

Identical by decent, 73
Inbreeding, 25, 73, 107, 145, 275
coefficient, 25, 148
depression , 26, 84, 147
lines, 147, 290
restricting, 177
Independent culling, 94
Indirect selection, 96
absorption of nutrient, 100
disease resistance, 97
fat deposition, 100
feed efficience,98
residual feed intake, 100
Inheritance, 18, 35
albinism, 19, 27
colour, 18
pigmentation, 23
polygenic, 35
scale pattern, 19, 20
Interaction, 191, 233, 244, 262
genotype-environment, 157, 233,

Kinship, 73

coefficient, 73, 76, 80

Least squares estimate, 225
Linkage map, 305
Locus, 21, 23, 57, 303

Maternal effects, 133, 129, 181
Mating design, 173, 178, 265
assortative, 128
balanced, 51
factorial, 179, 181, 265
full-sib, 25, 179
half-sib, 25, 280
hierarchical, 128, 265
nested, 180
random, 25, 74, 176
repeated, 245
single pairs, 177
self-fertilisation, 25
Maturation, 116, 258
Mean square (MS), 51,127
between classes, 52, 127
within classes, 52, 127
Meioses, 15, 288
Mendel's laws, 4, 17

Mendelian inheritance, 23, 35
Microsatellite, 264, 301, 302
Migration, 25, 31, 329
Mitoses, 15, 298
MME, 221
Model,
mixed, 130
multi-trait, 139
polygenic, 37, 41, 329
Monoculture, 236
Multiplier station, 193, 276, 281
Mutation, 14, 25, 28, 36, 84, 313

Nucleus, 9, 87
Nutrients, absorption, 100

Oogenese, 16
Overdominance, 152

Parameters, 45
statistical, 45

Path, 65, 76, 248
coefficient, 65
Pedigree, 77, 161, 251, 263, 267
Phenotype, 121
covariance, 121
parameters, 121
value, 197
Pleiotropy, 62
Polyculture, 236
Polymorphic, 21, 28, 36, 37, 302, 314
Population, 23, 197
base, 74, 93, 174, 330
closed, 74, 147
control, 243
effective size, 26, 161, 177, 321
genetics, 8, 23
idealized, 81
size, 26, 36, 186
synthetic, 154
wild, 109, 324, 330
Prediction of error, 213
Predictor, best, 199
Progeny test, 165, 188, 207, 262, 272
Protein, 14, 40
metabolic, 40
recipe, 12

QTL, 306, 312, 323
Quality, 258
colour, 259, 269
fat %, 259, 269
fat distribution, 259, 269
filet %, 259, 269
texture, 259, 270
Quantitative genetics, 8, 121, 251

Recapture frequency, 259
Recombination, 304, 307, 311
Regression, 59,
coefficient, 60, 125
Relationship, 59, 76, 160
between traits, 59
REML, 133, 248
Repeatability, 135
Response to selection, 102, 243
correlated, 96, 167, 274
cumulative, 94

Retension, 269
protein, 98
energy, 98
RNA, 10
MRNA, 14

Scale, 233
predicted value, 227
underlying, 51, 143, 169
Sea ranching, 259
Selection, 4, 27, 89, 322
artificial, 90, 322
BLUP, 185
broodstock, 273
coefficient, 27
combined, 165, 323
correlated response, 167
date of spawning, 105
differential, 91
directional, 42, 90, 321
disease resistance, 97, 102
divergent, 90, 245
domestication, 109, 322
family, 161, 205, 262, 271
feed conversion, 98, 105
full-sib, 204
half-sib, 205
index, 95,173, 201
individual, 159, 183, 191, 203
indirect, 96, 168
intensity, 27, 91, 166
limit, 106
marker assisted, 277, 312, 323
mass, 103, 159
methods, 159, 262
multiple trait, 95, 201
natural, 23, 84, 89, 107, 167, 244
proportion, 92
program, 8, 281
reciprocal-recurrent, 151
resistance, 7, 102
response, 91, 102, 164, 244
sexual maturation, 106
sib, 187, 204
stabilizing, 90, 321
tandem, 95
truncation, 86, 91, 101, 174, 200
within family, 85, 164, 185, 204
walkback, 85
Self-fertilization, 25
Sex, 113, 118, 283
chromosome, 114
maturation, 116, 238, 258, 295
ratio, 114, 183
reversal, 118, 290
SIP, 201, 217
SNP, 303
Socio-economic effect, 110, 265
Spawning, 179, 265, 272
Sperm, 14, 117, 291
Spermatogenesis, 16, 117, 295
Standard deviation, 38, 45, 108, 166
genetic, 39, 41
Stress, 3, 5

Tabular method, 78
Tags, 190, 257, 263
Tank effect, 228, 262
Test station, 281
Tetraploid, 287, 292, 298
Trait, 40, 47
Transformation of data, 140
Transgenic, 302, 315, 324
Threshold characters, 50, 140
Tribolium, 42, 108
Triploids, 287, 291, 296

Unisexual, 113

Variance, 45, 51,121
additive, 54, 121, 126, 176
analysis, 51, 127
component, 124, 129, 138
continuous, 50
dominance, 54, 121
effect of selection, 101
environmental, 55, 203
epistatic, 54, 121
genetic, 55, 178, 278
maternal, 57

Welfare, 5, 254

YY males, 297